

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: October 22, 2003, 12:19:30 : Search time 86 Seconds

(without alignments)
699,504 Million cell updates/sec

Title: US-09-805-311-2

Perfect score: 1939

Sequence: 1 MGICG1TKLADNAPKAKE.....SDKTSKAAANKKAGKRR 379

Scoring table: BLOSUM62

Gapop 10.0 , Capext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_15Jun03.*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*

2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*

3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*

4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*

5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*

6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*

7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*

8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*

9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*

10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*

11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*

12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*

13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*

14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*

15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*

16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*

17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*

18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*

19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*

20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*

21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*

22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1939	100.0	379	21	AAV95307
2	1939	100.0	379	21	AAV95309
3	1939	99.7	379	21	AAV95308
4	1933	99.7	379	21	AAV95310
5	1031	53.2	380	20	AAW92504
6	1007.5	52.0	377	20	AAW92505
7	998.5	51.5	385	22	ABB63960
8	996	51.4	378	20	AAW92508
9	938	48.4	382	20	AAW92505

10	699.5	36.1	373	22	ABG19545	Novel human data
11	648.5	33.4	340	18	AAW4216	Pyrococcus fur
12	648.5	33.4	340	19	AAW4216	Pyrococcus fur
13	648.5	33.4	340	19	AAW59440	Amino acid seq
14	648.5	33.4	340	23	ABG32941	Pyrococcus fur
15	647	33.4	326	19	AAW59953	Amino acid seq
16	644	33.2	332	22	AAW59951	Amino acid seq
17	636	32.8	343	22	AAW59951	Putative P. abyssi
18	632.5	32.6	343	20	AAW3778	Pyrococcus heat
19	609	31.4	325	19	AAW59950	Amino acid seq
20	600	30.9	340	19	AAW59949	Amino acid seq
21	562	29.0	340	19	AAW59952	Amino acid seq
22	560.5	28.9	326	18	AAW4215	Methanococcus j
23	560.5	28.9	326	19	AAW59939	Amino acid seq
24	560.5	28.9	326	23	ABG32942	Methanococcus j
25	553.5	28.5	326	19	AAW59969	Methanococcus j
26	538.5	27.8	336	19	AAW59982	Archaeoglobus
27	538.5	27.8	336	19	AAW59946	Amino acid seq
28	538.5	27.8	336	23	ABG32944	Archaeoglobus
29	532	27.4	328	19	AAW59948	Archaeoglobus
30	525	27.1	328	23	ABG32943	Archaeoglobus
31	525	27.1	328	23	ABG32943	Archaeoglobus
32	411	21.2	258	19	AAW59947	Methanobacteri
33	340.5	17.6	386	20	AAW42507	Amino acid seq
34	336	17.3	1516	21	AAW48195	Yeast delta-RAD
35	311	16.0	96	23	ABP3304	Plasmodium fal
36	257	13.3	726	22	ABW5192	Human nuclease
37	241.5	12.6	1236	22	ABW5297	Drosophila mel
38	241.5	12.6	1257	22	ABW5297	Drosophila mel
39	215.5	11.1	871	23	AAW4827	Thermophilic D
40	215.5	11.1	871	23	AAW4827	Thermophilic D
41	210	10.8	732	22	ABW7117	Drosophila mel
42	208	10.7	872	20	AAW31815	Thermococcus
43	205	10.6	872	18	AAW56605	Thermococcus
44	205	10.6	872	20	AAW56605	Thermococcus
45	205	10.6	872	20	AAW56605	Thermococcus

ALIGNMENTS

RESULT 1	AAV95307	standard: Protein: 379 AA.
ID	AAV95307	
XX	AAV95307	
AC	AAV95307	
XX		
DT	12-SEP-2000	(first entry)
XX		
DE	Maize Rad2/FEN-1 protein.	
XX		
KW	Maize; Rad2/FEN-1; transgenic plant; male sterile plant; endonuclease; exonuclease; DNA repair; gene targeting.	
XX		
OS	Zea mays.	
XX		
PN	W0200036109-A1.	
XX		
PD	22-JUN-2000.	
XX		
PF	16-NOV-1999; 99WO-US27147.	
XX		
FR	15-DEC-1998; 98US-0112332.	
XX		
FA	(PION-) PIONEER HI-BRED INT INC.	
PI	Mahajan Pa;	
XX		
LR	WPI: 2000-452026/39.	
DR	N-PSDB: AAA27923.	
XX		
PT	Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA recombination and repair in transgenic plants, e.g. for gene targeting.	

PT and the production of male sterile plants -
 XX
 XX Claim 11: Page 71-72: 85pp: English.
 XX
 CC The present sequence is that of maize RAD2/FEN-1, as deduced from
 CC a cDNA clone (see AAA27923) isolated from maize line B73 immature
 CC ear tissue. Rad2/FEN-1 is a structure specific endonuclease which
 CC under certain conditions also acts as an exonuclease. Rad2/FEN-1
 CC can be expressed in transgenic plant cells using conventional
 CC methods. The protein is involved in the regulation of DNA repair
 CC and recombination in plant systems and therefore may be used for
 CC improving gene targeting during further recombinant DNA protocols
 CC involving plants. RAD2/FEN-1 endonucleolytic activity is essential
 CC in DNA replication and nucleotide excision and repair reactions.
 CC The exolytic activity is involved in double strand break repair and
 CC end joining. The protein is also useful in strand exchange
 CC reactions during homologous recombination. These functions may be
 CC useful in gene targeting and in the production of male sterile
 CC plants. The efficacy of gene targeting can be improved by the
 CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can
 CC be produced by the down regulation of Rad2/FEN-1 expression.
 CC
 XX
 XX Sequence 379 AA:
 SQ
 Query Match 100.0%; Score 1939; DB 21; Length 379;
 Best Local Similarity 100.0%; Pred. No. 1,3e-170;
 Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGKGLTKLLADNAPKAKKEQKESYFGKRIADVASKSIYQFLIVGRGTMETLTNEAGE 60
 DB 1 MGKGLTKLLADNAPKAKKEQKESYFGKRIADVASKSIYQFLIVGRGTMETLTNEAGE 60
 QY 61 VTSHLQGMFNRTIRLLEAGIKPVYVFDGKPPDMKKOELAKRYSRDDATKDLTEAVSVD 120
 DB 61 VTSHLQGMFNRTIRLLEAGIKPVYVFDGKPPDMKKOELAKRYSRDDATKDLTEAVSVD 120
 QY 121 KDAIEKLSKRTYKVTROHNEDECKRLRLMGVYVVEASEAECAALCINDKVFVAASED 180
 DB 121 KDAIEKLSKRTYKVTROHNEDECKRLRLMGVYVVEASEAECAALCINDKVFVAASED 180
 QY 181 MDSLTFGAPRFLRHLMPPSSKKIPVMEFDVAKVLELELTMDQFIDLCITCGCDYCSIK 240
 DB 181 MDSLTFGAPRFLRHLMPPSSKKIPVMEFDVAKVLELELTMDQFIDLCITCGCDYCSIK 240
 QY 241 GIGGOTALKLIRHGSLESTLENLNKDRYQTPEDMPYQEARLKEFNVTLDPELKMTA 300
 DB 241 GIGGOTALKLIRHGSLESTLENLNKDRYQTPEDMPYQEARLKEFNVTLDPELKMTA 300
 QY 301 PDEGLISFLVKNQNGNEDRYTKAIEKISAKKSSQGRLESEFFKPTATTSAPLKRKETS 360
 DB 301 PDEGLISFLVKNQNGNEDRYTKAIEKISAKKSSQGRLESEFFKPTATTSAPLKRKETS 360
 QY 361 DKTSKAAANKKTKAGGKKK 379
 DB 361 DKTSKAAANKKTKAGGKKK 379
 RESULT 2
 ID AAY95309 standard; Protein: 379 AA.
 AC AAY95309:
 XX
 XX 12-SEP-2000 (first entry)
 DT
 XX
 XX Maize Rad2/FEN-1 protein.
 DE
 XX
 XX Maize: Rad2/FEN-1: transgenic plant; male sterile plant;
 KW endonuclease; exonuclease; DNA repair; gene targeting.
 XX
 XX Zea mays.
 OS
 XX
 PN W0200016109-A1.

XX
 PD 22-JUN-2000.
 XX
 XX 16-NOV-1999; 33MO-US27147.
 PF
 XX 15-DEC-1998; 380US-0112332.
 PR
 XX (P10N-) PIONEER HI-BRED INT INC.
 PA
 XX
 XX
 PI Mahajan PH:
 XX
 XX WPI: 2000-452026/39.
 DR
 XX N-PSDB; AAA27925.
 CR
 XX
 PT Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA
 PT recombination and repair in transgenic plants, e.g. for gene targeting
 PT and the production of male sterile plants -
 XX
 XX Example 1: Page 76-77: 85pp: English.
 PS
 CC The present sequence is that of maize RAD2/FEN-1, as deduced from
 CC a cDNA clone (see AAA27925) derived from maize line W23 tassel
 CC polyA RNA. Rad2/FEN-1 is a structure specific endonuclease which
 CC under certain conditions also acts as an exonuclease. Rad2/FEN-1
 CC can be expressed in transgenic plant cells using conventional
 CC methods. The protein is involved in the regulation of DNA repair
 CC and recombination in plant systems and therefore may be used for
 CC improving gene targeting during further recombinant DNA protocols
 CC involving plants. RAD2/FEN-1 endonucleolytic activity is essential
 CC in DNA replication and nucleotide excision and repair reactions.
 CC The exolytic activity is involved in double strand break repair and
 CC end joining. The protein is also useful in strand exchange
 CC reactions during homologous recombination. These functions may be
 CC useful in gene targeting and in the production of male sterile
 CC plants. The efficacy of gene targeting can be improved by the
 CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can
 CC be produced by the down regulation of Rad2/FEN-1 expression.
 CC
 XX
 XX Sequence 379 AA:
 SQ
 Query Match 100.0%; Score 1939; DB 21; Length 379;
 Best Local Similarity 100.0%; Pred. No. 1,3e-170;
 Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGKGLTKLLADNAPKAKKEQKESYFGKRIADVASKSIYQFLIVGRGTMETLTNEAGE 60
 DB 1 MGKGLTKLLADNAPKAKKEQKESYFGKRIADVASKSIYQFLIVGRGTMETLTNEAGE 60
 QY 61 VTSHLQGMFNRTIRLLEAGIKPVYVFDGKPPDMKKOELAKRYSRDDATKDLTEAVSVD 120
 DB 61 VTSHLQGMFNRTIRLLEAGIKPVYVFDGKPPDMKKOELAKRYSRDDATKDLTEAVSVD 120
 QY 121 KDAIEKLSKRTYKVTROHNEDECKRLRLMGVYVVEASEAECAALCINDKVFVAASED 180
 DB 121 KDAIEKLSKRTYKVTROHNEDECKRLRLMGVYVVEASEAECAALCINDKVFVAASED 180
 QY 181 MDSLTFGAPRFLRHLMPPSSKKIPVMEFDVAKVLELELTMDQFIDLCITCGCDYCSIK 240
 DB 181 MDSLTFGAPRFLRHLMPPSSKKIPVMEFDVAKVLELELTMDQFIDLCITCGCDYCSIK 240
 QY 241 GIGGOTALKLIRHGSLESTLENLNKDRYQTPEDMPYQEARLKEFNVTLDPELKMTA 300
 DB 241 GIGGOTALKLIRHGSLESTLENLNKDRYQTPEDMPYQEARLKEFNVTLDPELKMTA 300
 QY 301 PDEGLISFLVKNQNGNEDRYTKAIEKISAKKSSQGRLESEFFKPTATTSAPLKRKETS 360
 DB 301 PDEGLISFLVKNQNGNEDRYTKAIEKISAKKSSQGRLESEFFKPTATTSAPLKRKETS 360
 QY 361 DKTSKAAANKKTKAGGKKK 379
 DB 361 DKTSKAAANKKTKAGGKKK 379

Best local similarity 99.7%; Pred. No. 4,7e-170; Matches 378; Conservative 0; Mismatches 1; Indels 0; Gaps 0

QY 1 MGKIGLTLLADNAPKAKKEQKFESEYFGRIAVDASMSIYQFLIVGRISKFTLLINEAGS 60
 Jb 1 MGKIGLTLLADNAPKAKKEQKFESEYFGRIAVDASMSIYQFLIVGRISKFTLLINEAGS 60
 QY 61 VTSHLQGMFNRTIRLLFAGIKPVYVFQDKPPDKKQELARYSKRDATKDLTEAVEYSD 120
 Db 61 VTSHLQGMFNRTIRLLFAGIKPVYVFQDKPPDKKQELARYSKRDATKDLTEAVEYSD 120
 QY 121 KDAIEKLSKRTVKTROHNEQCKRLRLMGVYVFAESEAFAECALCINQXFFAVASED 160
 Db 121 KDAIEKLSKRTVKTROHNEQCKRLRLMGVYVFAESEAFAECALCINQXFFAVASED 160
 QY 181 MDLTFGAPRLRLHMDPSSKKIPVMEFDYAKYLELELMDQFIDLCILGCDYCSIR 240
 Db 181 MDLTFGAPRLRLHMDPSSKKIPVMEFDYAKYLELELMDQFIDLCILGCDYCSIR 240
 QY 241 GIGGOTALKLIRHOGSTESILENKKDRYOIPEDMPYQAFARLFEKPNVTLDIPELKWT 300
 Db 241 GIGGOTALKLIRHOGSTESILENKKDRYOIPEDMPYQAFARLFEKPNVTLDIPELKWT 300
 QY 301 PDEEGLISFLVKNDFNEDRVTKAIEKIKSAKNSSGRLSEFFKPTATTSAPLKRKETS 360
 Db 301 PDEEGLISFLVKNDFNEDRVTKAIEKIKSAKNSSGRLSEFFKPTATTSAPLKRKETS 360
 QY 361 DKTSAANKKTKAGGKK 379
 Db 361 DKTSAANKKTKAGGKK 379

RESULT 5
 AAM92504
 ID AAM92504 standard; Protein: 380 AA.

AC AAM92504:
 DT 23-APR-1999 (first entry)
 XX Human FEN-1 protein.
 DE Human FEN-1 protein.
 XX FEN-1; human; flap endonuclease; detection; diagnosis; carcinogen;
 KM neoplasma; antineoplastic agent; cleavage.
 XX Homo sapiens.
 OS Homo sapiens.
 PN J55874283-A.
 XX 23-APR-1999.
 PD 23-APR-1999.
 XX 30-MAY-1995; 95US-0455968.
 PF 30-MAY-1995; 95US-0455968.
 FR 30-MAY-1995; 95US-0455968.
 XX (HARR/) HARRINGTON J J.
 PA (HSIE/) HSIEH C.
 PA (LIEB/) LIEBER M R.
 XX Harrington JJ, Hsieh C, Lieber MR;
 PT WPI: 1999-179985/15.
 DR N-PSDB: AAX02107.
 XX DNA encoding flap endonuclease polypeptides - useful for producing
 PT e.g. recombinant polypeptides
 XX Claim 1; Fig 1A: 58pp; English.
 PS This sequence represents a human FEN-1 (flap endonuclease) protein. This
 CC protein can be used in methods for detecting a pathological condition in
 CC a patient, for diagnostic purposes, for screening for antineoplastic
 CC agents and carcinogens, for diagnostic staging of neoplasia, for

CC Producing recombinant flap endonuclease for use as research or
 CC diagnostic reagents, for producing antibodies reactive with the novel
 CC polypeptides, for producing transgenic nonhuman animals expressing the
 CC novel polypeptides encoded by a transgene. The invention also provides
 CC novel molecular cloning techniques and reagents involving cleavage of
 CC a flap or nick with a flap endonuclease.

QY 1 MGKIGLTLLADNAPKAKKEQKFESEYFGRIAVDASMSIYQFLIVGRISKFTLLINEAGS 60
 Db 1 MGKIGLTLLADNAPKAKKEQKFESEYFGRIAVDASMSIYQFLIVGRISKFTLLINEAGS 60
 QY 61 VTSHLQGMFNRTIRLLFAGIKPVYVFQDKPPDKKQELARYSKRDATKDLTEAVEYSD 120
 Db 61 VTSHLQGMFNRTIRLLFAGIKPVYVFQDKPPDKKQELARYSKRDATKDLTEAVEYSD 120
 QY 121 KDAIEKLSKRTVKTROHNEQCKRLRLMGVYVFAESEAFAECALCINQXFFAVASED 160
 Db 121 KDAIEKLSKRTVKTROHNEQCKRLRLMGVYVFAESEAFAECALCINQXFFAVASED 160
 QY 181 MDLTFGAPRLRLHMDPSSKKIPVMEFDYAKYLELELMDQFIDLCILGCDYCSIR 240
 Db 181 MDLTFGAPRLRLHMDPSSKKIPVMEFDYAKYLELELMDQFIDLCILGCDYCSIR 240
 QY 241 GIGGOTALKLIRHOGSTESILENKKDRYOIPEDMPYQAFARLFEKPNVTLDIPELKWT 300
 Db 241 GIGGOTALKLIRHOGSTESILENKKDRYOIPEDMPYQAFARLFEKPNVTLDIPELKWT 300
 QY 301 PDEEGLISFLVKNDFNEDRVTKAIEKIKSAKNSSGRLSEFFKPTATTSAPLKRKETS 360
 Db 301 PDEEGLISFLVKNDFNEDRVTKAIEKIKSAKNSSGRLSEFFKPTATTSAPLKRKETS 360
 QY 361 DKTSAANKKTKAGGKK 379
 Db 361 DKTSAANKKTKAGGKK 379

RESULT 6
 AAM92505
 ID AAM92505 standard; Protein: 377 AA.

AC AAM92505:
 DT 23-APR-1999 (first entry)
 XX Mouse FEN-1 protein.
 DE Mouse FEN-1 protein.
 XX FEN-1; mouse; flap endonuclease; detection; diagnosis; carcinogen;
 KM neoplasma; antineoplastic agent; cleavage.
 XX Mus sp.
 OS Mus sp.
 PN J55874283-A.
 XX 23-APR-1999.
 PD 23-APR-1999.
 XX 30-MAY-1995; 95US-0455968.
 PF 30-MAY-1995; 95US-0455968.
 FR 30-MAY-1995; 95US-0455968.
 XX (HARR/) HARRINGTON J J.
 PA (HSIE/) HSIEH C.
 PA (LIEB/) LIEBER M R.
 XX Harrington JJ, Hsieh C, Lieber MR;
 PT WPI: 1999-179985/15.
 DR N-PSDB: AAX02108.
 XX DNA encoding flap endonuclease polypeptides - useful for producing
 PT e.g. recombinant polypeptides
 XX Claim 1; Fig 1A: 58pp; English.
 PS This sequence represents a human FEN-1 (flap endonuclease) protein. This
 CC protein can be used in methods for detecting a pathological condition in
 CC a patient, for diagnostic purposes, for screening for antineoplastic
 CC agents and carcinogens, for diagnostic staging of neoplasia, for

XX DNA encoding flap endonuclease polypeptides - useful for producing
 PT e.g. recombinant polypeptides
 XX
 XX
 PS Claim 1: Fig 2A: 58pp; English.
 CC This sequence represents a mouse FEN-1 (flap endonuclease) protein. This
 CC protein can be used in methods for detecting a pathological condition in
 CC a patient, for diagnostic purposes, for screening for anti-neoplastic
 CC agents and carcinogens, for diagnostic staging of neoplasia, for
 CC producing recombinant flap endonuclease for use as research or
 CC diagnostic reagents, for producing antibodies reactive with the novel
 CC polypeptides, for producing transgenic nonhuman animals expressing the
 CC novel polypeptides encoded by a transgene. The invention also provides
 CC novel molecular cloning techniques and reagents involving cleavage of
 CC a flap or nick with a flap endonuclease.
 CC
 XX
 XX
 SQ Sequence 377 AA:
 Query Match 52.0%; Score 1007.5; DB 20; Length 377;
 Best Local Similarity 53.4%; Pred. No. 1,86-84;
 Matches 203; Conservative 68; Mismatches 100; Indels 9; Gaps 7;
 QY 1 MGIGLTKLADNAPKAKKEQFESYFGRIADVDSMSIYQFLIVGRTGMEITLNEAGE 60
 DB 1 MGIGLAKLADVAPSAIRNDIKSYFGRIADVDSMSIYQFLIVAV-ROGCDVILQNEDE 59
 QY 61 VSHLQGMENRTIRLENGIKPYVYFQCKPPDMKQELAKRYSKRDATKDLTEAVESGD 120
 DB 60 TTS-LMGMEYRTIR-MENGIRPYVYFQCKPPDLKSGELAKSERARAEKQIQQAQDEAG- 116
 QY 121 KDAIEKLSKRTVYKTRQHNEDCKRLRLMGVYVAVSEAEAECAALCINDKVAVASED 180
 DB 117 MEEVEKFTKRLVYKTRQHNEDCKRLRLMGVYVAVSEAEAECAALCINDKVAVASED 176
 QY 181 MDLTFGAPRFLRLHMDSSKKIPMEVDVAKVLELELTMDQFIDLCILCCDYCESIK 240
 DB 177 MDLTFGSPVLRHLTASEAKKLPIQEFHLSRVLOELTINQEOVDLCILCGSDYCESIR 236
 QY 241 GIGGCTALKLIRHGSISLELNKMDRQIPEDMPYOEARLFEKPNVT-LDIPELKMT 299
 DB 237 GIGAKRNVLLIOKHSISEIYRDLPSKYPVENMLHKEAQDLFEPEVPESEVEKMS 296
 QY 300 APDESGILSFLVKGDNEDRVTAKIEKISAKNKSOGRLSEFPKPAITSAFLKKET 359
 DB 297 EPNEBEIYKFMCGFQSFSEERIRSGVRLSKSRQGSTQGRIDDEFKVTGLSS-AKRKEP 355
 QY 360 SDKTSKAANKKTKAGGKK 379
 DB 356 EPKGS---AKKAKTGAGAK 372

RESULT 7
 ABB63960
 ID ABB63960 standard; Protein: 385 AA.
 AC ABB63960:
 XX
 XX 26-MAR-2002 (first entry)
 DE Drosophila melanogaster polypeptide SEQ ID NO 18672.
 XX
 XX Drosophila melanogaster polypeptide SEQ ID NO 18672.
 KM pharmaceutical.
 KW pharmaceutical.
 OS Drosophila melanogaster.
 OS
 PN WO200171042-A2.
 XX
 XX 27-SEP-2001.
 PD
 XX 23-MAR-2001; 2001WO-US09231.
 PE
 XX

PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 XX (PEKE) PE CORP NY.
 PA
 XX
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 XX WPI: 2001-656860/75.
 DR N-FSDB: ABI08063.
 XX
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 CC genes from Drosophila and for elucidating cell signalling and cell-cell
 CC interactions -
 CC
 XX
 XX
 PS Disclosure: SEQ ID NO 18672; 21pp + Sequence listing; English.
 CC
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA
 CC sequences (ABU1640-ABU16175) and the encoded proteins
 CC (ABU57737-ABU32072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIP
 CC at http://wipo.int/pub/published_pct_sequences.
 CC
 XX
 XX
 SQ Sequence 385 AA:
 Query Match 51.5%; Score 999.5; DB 22; Length 385;
 Best Local Similarity 52.9%; Pred. No. 1,30-83;
 Matches 202; Conservative 65; Mismatches 104; Indels 11; Gaps
 QY 1 MGIGLTKLADNAPKAKKEQFESYFGRIADVDSMSIYQFLIVGRTGMEITLNEAGE
 DB 1 MGIGLTKLADNAPKAKKEQFESYFGRIADVDSMSIYQFLIVGRTGMEITLNEAGE
 QY 61 VSHLQGMENRTIRLENGIKPYVYFQCKPPDMKQELAKRYSKRDATKDLTEAVESGD
 DB 60 TTS-LMGMEYRTIR-LDNGIKPYVYFQCKPPDLKSGELAKSERARAEKQIQQAQDEAG-
 QY 121 KDAIEKLSKRTVYKTRQHNEDCKRLRLMGVYVAVSEAEAECAALCINDKVAVASED
 DB 120 DAGIEKFNRLVYKTRQHNEDCKRLRLMGVYVAVSEAEAECAALCINDKVAVASED
 QY 181 MDLTFGAPRFLRLHMDSSKKIPMEVDVAKVLELELTMDQFIDLCILCCDYCESIK
 DB 180 MDLTFGAPRFLRLHMDSSKKIPMEVDVAKVLELELTMDQFIDLCILCCDYCESIK
 QY 241 GIGGCTALKLIRHGSISLELNKMDRQIPEDMPYOEARLFEKPNVT-LDIPELKMT
 DB 240 GIGPKRAIYLINTYRIETLIDNLSSKYTVENMYKARELFEPEVADSDILKAV
 QY 300 APDESGILSFLVKGDNEDRVTAKIEKISAKNKSOGRLSEFPKPAITSAFLKKET
 DB 300 EPDESGILSFLVKGDNEDRVTAKIEKISAKNKSOGRLSEFPKPAITSAFLKKET
 QY 355 KKEKTSKTSKAANKKTKAGG 375
 DB 358 KKEKTSKTSKAANKKTKAGG 377

RESULT 8
 AAW92508
 ID AAW92508 standard; Protein: 378 AA.
 AC AAW92508:
 XX
 XX
 XX 23-APR-1999 (first entry)
 DE Human FEN-1 protein.
 XX

KW FEN-1: human; flap endonuclease; detection; diagnosis; carcinogen;
 KM neoplasma; antineoplastic agent; cleavage.
 XX Homo sapiens.
 OS
 PN US5874283-A.
 XX
 PD 23-FEB-1999.
 XX
 PF 30-MAY-1995: 55US-0455968.
 XX
 PR 30-MAY-1995: 55US-0455968.
 XX
 PA (HARR/) HARRINGTON J J.
 PA (HSIE/) HSIEH C.
 PA (LIEBER/) LIEBER M R.
 XX
 PI Harrington JJ, Hsieh C, Lieber MR;
 XX
 DR WPI: 1999-179985/15.
 DR N-PSDB: AAX02111.
 XX
 PT DNA encoding flap endonuclease polypeptides - useful for producing
 PT e.g. recombinant polypeptides
 XX
 PS Disclosure: Fig 5A-B: 58pp: English.
 XX
 CC This sequence represents a human FEN-1 (flap endonuclease) protein. This
 CC protein can be used in methods for detecting a pathological condition in
 CC a patient, for diagnostic purposes, for screening for antineoplastic
 CC agents and carcinogens, for diagnostic staging of neoplasia, for
 CC producing recombinant flap endonuclease for use as research or
 CC diagnostic reagents, for producing transgenic nonhuman animals expressing the
 CC polypeptides, for producing antibodies reactive with the novel
 CC novel polypeptides encoded by a transgene. The invention also provides
 CC novel molecular cloning techniques and reagents involving cleavage of
 CC a flap or nick with a flap endonuclease.
 CC
 XX
 SO Sequence 378 AA:
 51.4% Score 996; DB 20; Length 378;
 Query Match Best Local Similarity 52.4%; Pred. No. 2, 1e-83;
 Matches 199; Conservative 70; Mismatches 103; Indels 8; Gaps 6;
 QY 1 MGKGLTKLLADNAPAKMEQKESYFGKRIAVDASMSIYQFLIVGRTGMETLTNEAGE 60
 DB 1 MEIHGLAKLLADYAPSAIRENDIKSYGRKVAIDASMSIYQFLIV-ROGDVLTQREFGE 59
 QY 61 VTSHLOGMNRITRLLEAGIKPYVFDGKPPDKKQELAKKRYSKRDAATKDLTEAVEYGD 120
 DB 60 TTS-LMGMFRTKMK-MENGIKPYVFDGKPPQKSGPLAKRSERRAEKQLOQAQAGM 117
 QY 121 KDAIEKLSKRTVAVTROHNECKRLRLMGVYVFAVSEAEACALCINDKVFVAASED 180
 DB 118 EEEVEKFTKRLVAVTKOHNECKHLSTMGIPYLDAPSEAEACALAAKAKVVAATED 177
 QY 181 MDLITGAPRFLKHLMDPSSKIPVMEFVYAKVLELELTMDQFDICITCGDYKDSKR 240
 DB 178 MDCLTFGSPVLMKHLVASEAKKLPIDQFHLISVLADELGNQEFVDCITLQSDYKDSKR 237
 QY 241 GIGCOTAKLTIHQSGSTILENKKDRVOIPEDWYVCFARLCKEPNPT-LDIPFLAKT 299
 DB 238 GIGAKKAVNDLICHHKSTEEVYRLIDPSKYVPENNMLHKEKQDLEFPEVVDSEVLAKS 297
 QY 350 APCRFLISFLVNDNGENRVTAKIKKISAKNKKSSQGLRESFFPKATTSAPKPKST 359
 DB 299 EPNFEELVAKMCEKQOFFEERIRSGVKRLSKPSQSTQGLDQDFKVTSLSS-AKRKEP 356
 QY 350 SDKTSKAAAKKTKAGCKKK 379
 DB 357 E---PKGPAPKAKKATGAGCK 373

RESULT 9
 ID AAM92506
 XX AAM92506 standard; Protein: 382 AA.
 AC
 XX AAM92506:
 XX
 DT 23-APR-1999 (first entry)
 XX
 DE Yeast FEN-1 protein.
 XX
 KM FEN-1: yeast; flap endonuclease; detection; diagnosis; carcinogen;
 KM neoplasma; antineoplastic agent; cleavage.
 XX
 OS Saccharomyces cerevisiae.
 PN US5874283-A.
 XX
 PD 23-FEB-1999.
 XX
 PF 30-MAY-1995: 55US-0455968.
 XX
 PR 30-MAY-1995: 55US-0455968.
 XX
 PA (HARR/) HARRINGTON J J.
 PA (HSIE/) HSIEH C.
 PA (LIEBER/) LIEBER M R.
 XX
 PI Harrington JJ, Hsieh C, Lieber MR;
 XX
 DR WPI: 1999-179985/15.
 DR N-PSDB: AAX02109.
 XX
 PT DNA encoding flap endonuclease polypeptides - useful for producing
 PT e.g. recombinant polypeptides
 XX
 PS Disclosure: Fig 3A: 58pp: English.
 XX
 CC This sequence represents a yeast FEN-1 (flap endonuclease) protein. This
 CC protein is used in a method to isolate novel human FEN-1 proteins for
 CC detecting a pathological condition in a patient, for diagnostic purposes
 CC for screening for antineoplastic agents and carcinogens, for diagnostic
 CC staging of neoplasia, for producing recombinant flap endonuclease for u
 CC as research or diagnostic reagents, for producing antibodies reactive
 CC with the novel polypeptides, for producing transgenic nonhuman animals
 CC expressing the novel polypeptides encoded by a transgene. The invention
 CC also provides novel molecular cloning techniques and reagents involving
 CC cleavage of a flap or nick with a flap endonuclease.
 CC
 XX
 SO Sequence 382 AA:
 48.4% Score 938; DB 20; Length 382;
 Query Match Best Local Similarity 49.5%; Pred. No. 4, 9e-78;
 Matches 190; Conservative 69; Mismatches 109; Indels 16; Gaps
 QY 1 MSIKGLTKLLADNAPAKMEQKESYFGKRIAVDASMSIYQFLIVGRTGMETLTNEAGE 6
 DB 1 MSIKGLTKLLADNAPAKMEQKESYFGKRIAVDASMSIYQFLIVGRTGMETLTNEAGE 6
 QY 61 VTSHLOGMNRITRLLEAGIKPYVFDGKPPDKKQELAKKRYSKRDAATKDLTEAVEYGD 1
 DB 61 TSHLMGAFVRLTALMDGKIKPYVFDGKPPQKSGPLAKRSERRAEKQLOQAQAGM 1
 QY 121 KDAIEKLSKRTVAVTROHNECKRLRLMGVYVFAVSEAEACALCINDKVFVAASED 1
 DB 121 KCAIEKLSKRTVAVTROHNECKRLRLMGVYVFAVSEAEACALCINDKVFVAASED 1
 QY 221 K---MKCEPRLVAKSKENHFEAKVIGLIGIPYIILAFLAQAQAFKAKKGVAAASED 1
 DB 221 K---MKCEPRLVAKSKENHFEAKVIGLIGIPYIILAFLAQAQAFKAKKGVAAASED 1
 QY 241 GIGCOTAKLTIHQSGSTILENKKDRVOIPEDWYVCFARLCKEPNPT-LDIPFLAKT 2
 DB 241 GIGCOTAKLTIHQSGSTILENKKDRVOIPEDWYVCFARLCKEPNPT-LDIPFLAKT 2
 QY 350 SDKTSKAAAKKTKAGCKKK 379
 DB 357 E---PKGPAPKAKKATGAGCK 373

QY 256 --LKWAPDEEGLSFLVKNQGNEDRVTKAIEKISAKNKSOGRIEFPKPAIT--- 350
DQ 237 INLKSIPKREKELIYLDDOKKFSERKYSGLSLKKGKSGJGGRLDGFGVYKKEKQ 356
QY 351 -SAPLKRETSDDKTSKAAANKKTK 373
DB 357 LAAAKRAQENKKLKN-KNKKVK 375
RESULT 10
ID ABG19545 standard; protein: 373 AA.
AC ABG19545:
XX
XX 13-FEB-2002 (first entry)
DE Novel human diagnostic protein #.9536.
XX
XX Human: chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
OS
XX MO200275067-A2.
XX
XX 11-OCT-2001.
PD
XX 30-MAR-2001; 2001WO-0508631.
PF
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0549167.
XX
XX (Hvsh-) HVSEQ INC.
PA
XX
PI Dmanac RT, Liu C, Tang YF;
DR WPI: 2001-639362/73.
DR N-PSDB: AAS83732.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
PS
PS Claim 20: SEQ ID NO 49904; 103bp; English.
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPD
CC at ftp.wipd.int/pub/published_pcl_sequences.
XX
XX Sequence 373 AA:
Query Match 36.1%; Score 659.5; DB 22; Length: 373;

Best Local Similarity 41.1%; Pired. No. 5,4e-5%;
Matches 155; Conservative 68; Mismatches 105; Indels 49; Gaps
QY 1 KGIKGLTKLLADNAPRAMKQKKESTFGRRIAVASMSTYCLILVGRQMETLINEAGE
DB 35 MGIOGLAKLLADVAPSAIENDIKSYGRRVAIDASMSIYHFLIAV-ROGGDVLONEGE
QY 61 VTSHLQGMENRTIRLLEAGIKPYVFDGKIPDOKKQELAKRYSKRDQATKDLTEAVEGD
DB 95 TSHLMKQMYRTIRKRNCKIKPLY-----KSGELAKCSFLRAFAFKOLOQAAQAA
QY 121 KDAIEKLSKRTVAVTTRQHNEDCKRLJLMDGPVVEAASEAARCAALCINDKFAVASED
DB 146 EGYVEKTKVLAVVATKQHNCKCHLLSLMGIPLDAPIAFAACALVAKGKYPAATED
QY 181 MDSLTSGAPRFLNHLMDPSSKILPWFEDVAKYLELELIMDOFIDLCILGCTCYDSIK
DB 206 MDCFTGSPVLMQHLTASS-----DYCKSIW
QY 241 GIGGOTALKLIRQHSFISILEMLNDRQIPEDMRYOEARLRFKPNVT-LDIPELKWT
DB 232 SIGKRAVDILQHKSTEEVQMLDPRKPYLPNNRLHKEAYCFLFEFV-GPFSVILKMS
QY 300 APPEEGLSIFLVKNQGNEDRVTKAIEKISAKNKSOGRIEFPKPAITSAFLKRE-
DB 292 EPNEDELVKFVCGEKEQFSEERIRHSYVKRLKTEKG-GGEFGPGWQIP--GRASPIKRPDH
QY 359 TSDKTSKAAANKKTKAG 375
DB 349 QEDQHTSRSSDSGSHKG 365
RESULT 11
ID AAN24216 standard; protein: 340 AA.
AC AAN24216:
XX
XX 25-MAR-2003 (updated)
DE 14-APR-1998 (first entry)
XX
XX Pyrococcus furiosus FEN-1 endonuclease.
XX Nucleic acid cleavage; 5' nuclease; DNA cleavage; RNA cleavage;
XX invader directed cleavage; FEN-1: endonuclease.
XX
XX Pyrococcus furiosus.
XX
XX MO9727214-A1.
FN
XX
XX 31-JUL-1997.
PD
XX
XX 22-JAN-1997; 97WO-0501072.
PF
XX
XX 02-DEC-1996; 96US-0759038.
PR 24-JAN-1996; 96US-0599491.
PR 12-JUL-1996; 96US-0682853.
PR 29-NOV-1996; 96US-0756386.
PR 02-DEC-1996; 96US-0758314.
XX
XX (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX
XX Hall JG, Lyamichev VI, Prudent JR, Brow MWD, Kaiser MW;
PI Lyamichev N, Olive DM, Dahlberg JE;
DB WPI: 1997-39361/36.
DB N-PSDB: AAT76685.
XX
XX Thermostable structure-specific nuclease(s) - used for detection and
PI characterisation of nucleic acid sequences and variations in nucleic
PT acid sequences
XX
XX Example 28; Page 283-285; 457pp; English.

XX This sequence comprises *Pyrococcus furiosus* (Pfu) FEN-1
 CC endonuclease. Large-scale production of the enzyme was performed
 CC using *E. coli* host cells transformed with a vector carrying the Pfu
 CC FEN-1 coding sequence (see AAT76685). Pfu FEN-1 is a thermostable
 CC enzyme. It can be used in novel methods for the detection and
 CC characterisation of nucleic acid sequences and variations in
 CC nucleic acid sequences.
 CC (updated on 25-MAR-2003 to correct PI field.)
 CC
 XX Sequence 340 AA;
 SQ
 Query Match 33.4%; Score 648.5; DB 18; Length 340;
 Best Local Similarity 41.3%; Pred. No. 2.4e-51;
 Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;
 QY 19 KEQKFESYFGKRAVDASMSITVGTGEMETLTNEGVSHTLQGMNRTLLLEA 79
 DB 12 KEIELENLYGKRIADALNAIYQFLSTIRKQDCTPLMSKGRITSHLSGLFYKINLMEA 71
 QY 79 GIKPYVYFDGPPDKKQELAKRSKRDATKDLTEAVGVGDKDAIEKLSKRYKTAGH 138
 DB 72 GIKPYVYFDGPPDKKQELAKRSKRDATKDLTEAVGVGDKDAIEKLSKRYKTAGH 138
 QY 139 NEDCKRLRLMGVYVPAVSEAECAALCINDKFAVASEDMOSTFGAPRLRLHMDP 198
 DB 132 IEDAKKLLLELMGIPYVQAPSEGEAOAAVMKAGSYVASODYDLSLFGAPRLVRLNLTIT 191
 QY 199 SSKKIP-----VMEFDVAKYLELELTMDQFIDLCIGCDY-CDSIKGIG 244
 DB 192 GKRKLPGKKNVYEIKPELITLLE-----EVKLKLTREKTLALVAGTDINPGIKIGI 247
 QY 245 QTALKLIRHGSISTELENKND---RYQIPEDMPYOEARLRFKPNVTIDIPELKWTAP 301
 DB 248 KKALEIVRH-----SKDPLAKFKOSVDLYAIKEFPLNPVT-DNYNLVWROP 295
 QY 302 DEEGLISFLVKNQGNEDVYTKALEIKSAKNSKSGSTLESFVK 345
 DB 296 DEEGLKFLCDHDFSEERVKNGLERLKAIKSGKSTLESFVK 339
 RESULT 12
 ID AAM79970 standard: Protein: 340 AA.
 XX
 AC AAM79970:
 XX
 DT 02-FEB-1999 (first entry)
 XX
 DE *Pyrococcus furiosus* FEN-1 endonuclease.
 XX
 KM Nucleic acid detection; multiple sequential invasive cleavage;
 XX FEN-1; endonuclease; nuclease.
 OS *Pyrococcus furiosus*.
 XX
 PN W09842873-A1.
 XX
 ED 01-OCT-1998.
 XX
 EF 24-MAR-1998; 54WD-GS05809.
 XX
 PR 24-MAR-1997; 97US-0823516.
 XX
 PA (THIRD) THIRD WAVE TECHNOLOGIES INC.
 XX
 PI Brock MAC, Hall JG, Kwiatkowski RW, Lyamichov V;
 PI Mast AL, Vavra SH;
 XX WPI: 1998-557036/47.
 CR N-PSDB; AAV6840.
 XX
 PT Detecting target nucleic acid by sequence-specific cleavage of

PT complex with two specific oligonucleotides - used to detect
 PT cytomegalovirus DNA
 XX
 PS Example 28b: Page 316-317; 52app; English.
 XX
 CC This is the amino acid sequence of FEN-1 endonuclease of
 CC *Pyrococcus furiosus* (Pfu). FEN-1 DNA (see AAM65840) has been
 CC ligated into vector for expression in *E. coli* cells. The invention
 CC relates to means for the detection and characterisation of nucleic
 CC acid sequences, and variations in nucleic acid sequences. It also
 CC relates to methods for forming a nucleic acid cleavage structure on
 CC a target sequence and cleaving this structure in a site-specific
 CC manner, preferably using a thermostable structure-specific nuclease
 CC such as FEN-1. Cleavage of the cleavage structure by the
 CC nuclease indicates the presence of specific nucleic acid sequences
 CC or specific variants. The invention further relates to methods for
 CC the separation of nucleic acid molecules based on charge, methods
 CC for the detection of non-target cleavage products via the formation
 CC of a complete and activated protein binding region, and methods for
 CC the detection of nucleic acid from various viruses (e.g. human
 CC cytomegalovirus) in a sample. The method amplifies the detection
 CC molecule rather than the target itself, is less subject to
 CC contamination than exponential amplification processes, and allows
 CC many targets to be analysed in a single reaction.
 CC
 XX Sequence 340 AA;
 SQ
 Query Match 33.4%; Score 648.5; DB 19; Length 340;
 Best Local Similarity 41.3%; Pred. No. 2.4e-51;
 Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;
 QY 19 KEQKFESYFGKRIAVDASMSITVGTGEMETLTNEGVSHTLQGMNRTLLLEA 79
 DB 12 KEIELENLYGKRIADALNAIYQFLSTIRKQDCTPLMSKGRITSHLSGLFYKINLMEA 71
 QY 79 GIKPYVYFDGPPDKKQELAKRSKRDATKDLTEAVGVGDKDAIEKLSKRYKTAGH 138
 DB 72 GIKPYVYFDGPPDKKQELAKRSKRDATKDLTEAVGVGDKDAIEKLSKRYKTAGH 138
 QY 139 NEDCKRLRLMGVYVPAVSEAECAALCINDKFAVASEDMOSTFGAPRLRLHMDP 198
 DB 132 IEDAKKLLLELMGIPYVQAPSEGEAOAAVMKAGSYVASODYDLSLFGAPRLVRLNLTIT 191
 QY 199 SSKKIP-----VMEFDVAKYLELELTMDQFIDLCIGCDY-CDSIKGIG 244
 DB 192 GKRKLPGKKNVYEIKPELITLLE-----EVKLKLTREKTLALVAGTDINPGIKIGI 247
 QY 245 QTALKLIRHGSISTELENKND---RYQIPEDMPYOEARLRFKPNVTIDIPELKWTAP 301
 DB 248 KKALEIVRH-----SKDPLAKFKOSVDLYAIKEFPLNPVT-DNYNLVWROP 295
 QY 302 DEEGLISFLVKNQGNEDVYTKALEIKSAKNSKSGSTLESFVK 345
 DB 296 DEEGLKFLCDHDFSEERVKNGLERLKAIKSGKSTLESFVK 339
 RESULT 12
 ID AAM59540 standard: Protein: 340 AA.
 XX
 AC AAM59540:
 XX
 DT 21-DEC-1998 (first entry)
 XX
 DE Amino acid sequence of the Pfu FEN-1 endonuclease.
 XX
 KM Mja FEN-1 endonuclease; Tag gene; structure-specific nuclease;
 KM mutant; DNA polymerase; bacterial; fungi; protozoa; RNA virus;
 XX hepatitis C virus; HIV; thermostable.
 OS *Pyrococcus furiosus*.
 XX
 PN W09821774-A1.


```

XX 04-JUN-1998.
XX
XX 26-NOV-1997: 97WD-US21783.
XX
XX 02-DEC-1996: 96US-0758314.
XX
XX 29-NOV-1996: 96US-0757653.
XX
XX (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX
XX Kaiser MW, Lyamichev V.I., Lyamicheva N:
XX MPI: 1998-322748/28.
XX
XX N-PSDB: AAV53951.
XX
XX Thermostable structure-specific nuclease(s) derived from mutant DNA
XX polymerase(s) - useful for detecting mutant allele(s) or strains of
XX microorganisms
XX
XX Example 29: Pages 280-281; 472pp; English.
XX
XX This is the amino acid sequence encoding the Pfu FEN-1 endonuclease,
XX used in the method of the invention. In this process thermostable
XX structure-specific nucleases are derived from mutant DNA polymerases,
XX which can be used for detecting mutant alleles or strains of
XX microorganisms. The structure-specific nucleases can be used in
XX mixtures, compositions and kits to treat nucleic acid, e.g. for
XX detection of wild type and mutant alleles of genes, for detection
XX and/or identification of strains of microorganisms such as bacteria,
XX fungi, protozoa, especially for detection of RNA viruses such as the
XX hepatitis C virus (HCV).
XX
XX Sequence 340 AA:
XX
XX Query Match 33.4%; Score 648.5; DB 19; Length 340;
XX Best Local Similarity 41.3%; Pred. No. 2.4e-51;
XX Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;
XX
XX 19 KEKFESEYGRKIAVDASMS:YQFLIVGRTGKMETLTNEAGEVTS:QGMENRTICLLEA 78
XX 12 KEIENLNLGKKIADLNAL:YQFLSTIRQKDGITPLMSKGRITSHLSGLFRTINLMEA 71
XX
XX 79 GIKPVYVFDGKPPDKKKQQLAKRYSKRDDATKDLTEAVEVGDKDA:EKLSKRTYKVTROH 138
XX 72 GIKPVYVFDGKPPDKKKQQLAKRYSKRDDATKDLTEAVEVGDKDA:EKLSKRTYKVTROH 131
XX 139 NEDCKRLRLMGVPPVPAPEAEACALCTNDKVFVASEDMSTTGAPRLRLH:MEP 198
XX 132 IEDAKRLLELWGIPI:VOAPSEGAQAAVMAAKGSYVASODYSOSLLFGAPRLVRLN:TTT 191
XX 199 SSKRKP-----VMEFVAKVLELELTMDQFIDLCITGCDY-CDSITKGIIG 244
XX 192 GKRLKPKRKNVYEIRPELTILE-----EVLKELKLTREKLTIELALVGTIDYNDGSKGIC 247
XX 245 QTAALKLIRQHGSIESTLENLND---RYQIPEDWPYQEARLREKPNVTLIDIPRLKWTAP 301
XX 248 KKALEIVRH-----SKDPLAKRQKQSDVDLVAIKKEFLNPVPT-DNYNLVWRDP 295
XX
XX 302 DEGLISFLVKNDFNEDRVTKAIEKIKSAKNSKSGRLSEFFK 345
XX 296 DEGLIKFLCDHDFSEERVKNGLERLKAIRKATKSCNSTLESFVK 339
XX
XX RESULT 14
XX ABG92941
XX ID ABG92941 standard: Protein; 340 AA.
XX
XX AC ABG92941:
XX
XX 20-NOV-2002 (first entry)
XX
XX Pyrococcus furiosus FEN1 protein.
XX

```

```

XX Phosphoramidite: INVADER assay cleavage reaction; FEN1: cleavage;
XX nucleic acid separation; DNA polymerase: human; MCF-1: unglycylated;
XX monocytic chemottractant protein-1: enzyme.
XX
XX Pyrococcus furiosus.
XX
XX W0200263040-A2.
XX
XX 15-ACG-2002.
XX
XX 06-FEB-2002: 2002WD-US03423.
XX
XX 06-FEB-2001: 2001US-0777430.
XX
XX (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX
XX Lyamichev V., Skrzeczynski Z., Allawi HT, Mayland SR, Yakova T;
XX Neri BP:
XX MPI: 2002-674850/72.
XX
XX Composition useful for e.g. separation of nucleic acids comprises a
XX positively or neutrally charged phosphoramidite -
XX
XX Example 2: Page 179-180; 197pp; English.
XX
XX The invention relates to a composition comprising a positively or
XX neutrally charged phosphoramidite. The composition is useful for
XX separation of nucleic acid molecules. The composition is further useful
XX for fractionation of specific nucleic acids by selective charge revers
XX useful in e.g. INVADER assay cleavage reactions) and in the synthesis
XX of charge-balanced molecules. In the fractionation of nucleic acid
XX molecules, the method provides an absolute readout of the partition of
XX products from substrates (i.e. provides a 100% separation). Through th
XX use of multiple positively charged adducts, synthetic molecules can be
XX constructed with sufficient solubilization due to the fact that the
XX normally negatively charged strand is made nearly neutral. It is also
XX possible to distinguish between an enzymatically or thermally degraded
XX DNA fragments due to the absence or presence of 3' phosphate.
XX
XX ABG92335 ABG92945 represent amino acid sequences of the invention.
XX
XX Sequence 340 AA:
XX
XX Query Match 33.4%; Score 648.5; DB 23; Length 340;
XX Best Local Similarity 41.3%; Pred. No. 2.4e-51;
XX Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;
XX
XX 19 KEKFESEYGRKIAVDASMS:YQFLIVGRTGKMETLTNEAGEVTS:QGMENRTICLLEA 78
XX 12 KEIENLNLGKKIADLNAL:YQFLSTIRQKDGITPLMSKGRITSHLSGLFRTINLMEA 71
XX
XX 79 GIKPVYVFDGKPPDKKKQQLAKRYSKRDDATKDLTEAVEVGDKDA:EKLSKRTYKVTROH 138
XX 72 GIKPVYVFDGKPPDKKKQQLAKRYSKRDDATKDLTEAVEVGDKDA:EKLSKRTYKVTROH 131
XX 139 NEDCKRLRLMGVPPVPAPEAEACALCTNDKVFVASEDMSTTGAPRLRLH:MEP 198
XX 132 IEDAKRLLELWGIPI:VOAPSEGAQAAVMAAKGSYVASODYSOSLLFGAPRLVRLN:TTT 191
XX 199 SSKRKP-----VMEFVAKVLELELTMDQFIDLCITGCDY-CDSITGIGIG 244
XX 192 GKRLKPKRKNVYEIRPELTILE-----EVLKELKLTREKLTIELALVGTIDYNDGSKGIC 247
XX 245 QTAALKLIRQHGSIESTLENLND---RYQIPEDWPYQEARLREKPNVTLIDIPRLKWTAP 301
XX 248 KKALEIVRH-----SKDPLAKRQKQSDVDLVAIKKEFLNPVPT-DNYNLVWRDP 295
XX
XX 302 DEGLISFLVKNDFNEDRVTKAIEKIKSAKNSKSGRLSEFFK 345
XX 296 DEGLIKFLCDHDFSEERVKNGLERLKAIRKATKSCNSTLESFVK 339
XX
XX RESULT 15
XX

```


R:Harrington, J.J.; Lieber, M.R.

Genes Dev. 8, 1344-1355, 1994

A>Title: Functional domains within FEN-1 and RAD2 define a family of structure-specific

A:Reference number: A53730; M01D:95011546; PMID:7926735

A:Accession: A53730

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-378 <RAR>

A:Cross-references: GB:126320; NID:9499648; P1DN:CAA37664.1; P1D:9499649

C:Superfamily: conserved hypothetical protein YK113C

C:Keywords: DNA repair; endonuclease

Query Match 52.3%; Score 1014; DB 2; Length 378;
Best Local Similarity 53.2%; Pred. No. 1,66-60;
Matches 202; Conservative 69; Mismatches 101; Indels 8; Gaps 6;

QY 1 MGKIGITKLADNAAKAKKEQFESEYFGKRIAVDASMSIYQFLIVYGRTGHTLTNEAGE 60
DB 1 MGHIQIAKLIADVASAIKRENDIKSYFGRKVAIDASMSIYQFLIIV-ROGSDVLNENGE 59
QY 61 VISHLQGMENRTIRLEAGIKPYVYEDGKPPDMKQOELAKRYSKRDATXDLTAVENGD 120
DB 60 TTS-LMCMFYRTIR-MENKIKPYVYFDGKPPOLAKSGELAKSRRAEKEKQLOOQOEGM 117
QY 121 KDAIEKLSKRTVYKTRQHNEDCKRLRLMGVPEVPEAPSEAEACALCINDVFAVASED 180
DB 118 EREVERKTRIVKVKYKOHNEDECKHLISMGHYLDAPSEAEACALAKAGVVAATED 177
QY 181 MDSITFGAPRFLRHLMDSKSKIPVPEFDVAKVLELTLMDQFLDLCAGGQYCSIR 240
DB 178 MDCLIFGSPVLMRLHASEAKKLPLOERHLSKVLDELINOCFVDLCLIGSDYESTIR 237
QY 241 GIGGQATALKLRQSSIESILENLNKDRYOIPEDFYQEARLFEKPNVT-ID-PELKMT 299
DB 236 GIGAKRAVDLIQKHSEIEVIRLDPSKYVPDENLHLEAQQLFEZVVCPESELKMS 297
QY 300 APDEGLISFLYKDNQFNEDRYTKAIEKISAKKNSOGRLSPFKPATTSAPLKSKET 359
DB 298 EPNEELVFMCGEKFSEIRISGVKRLSKRSRGSTGRIDDFPKYVGSLS-AKREKP 356
QY 350 SPKTSKAANKKTKAGKKK 375
DB 357 E---FKGPAKKKAKTKGCGAK 373

RESULT 5
113692

hypothetical protein ECG003.3 - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000

C:Accession: T13692

R:Karpay, L.; Harris, D.; Batrcell, R.

submitted to the EMBL Data Library, November 1998

A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.

A:Reference number: Z17699

A:Accession: T13692

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-385 <MUR>

A:Cross-references: EMBL:AL031863; P1DN:CAA21320.1

C:Genetics:

A:Cross-references: FlyBase:FBgn0025833

A:Introns: 28/1

A:Note: ECG003.3

C:Superfamily: conserved hypothetical protein YK113C

Query Match 51.5%; Score 998.5; DB 2; Length 385;
Best Local Similarity 52.9%; Pred. No. 1,86-59;
Matches 202; Conservative 65; Mismatches 104; Indels 11; Gaps 5;

QY 1 MGKIGITKLADNAAKAKKEQFESEYFGKRIAVDASMSIYQFLIVYGRTGHTLTNEAGE 60
DB 1 MGITLSKLIADLAPQAIRSEMKHFFGRKVAIDASMSIYQFLIIVRSEGAQ-LATVMD 59

QY 61 VISHLQGMENRTIRLEAGIKPYVYEDGKPPDMKQOELAKRYSKRDATKDLTAVENGD
DB 60 PISHLMGMEYRTIRLDNDGKIPVYEDGKPPDLKSOELAKRERREAEKALKAAATDAD
QY 121 KDAIEKLSKRTVYKTRQHNEDCKRLRLMGVPEVPEAPSEAEACALCINDVFAVASED
DB 120 DACIEKFNRLVAVITKEHAKKEKELLTMDQVYVDAPCEAEACALVAGKRYATATED
QY 181 MDSITFGAPRFLRHLMDSKSKIPVPEFDVAKVLELTLMDQFLDCLTGGDVCDSIR
DB 180 MDALTYGSLTLLRYLTYSKAKKPVAFKFDKLLDEGLAIINREFFIDCLTGGDVCDSIR
QY 241 GIGGQATALKLRQSSIESILENLNKDRYOIPEDFYQEARLFEKPNVT-ID-PELKMT
DB 240 GIGKRAELILINRYRDETLLDMLDSKRTVPEPMNKVYARLFIPEVADSDILKAV
QY 300 APDEGLISFLYKDNQFNEDRYTKAIEKISAKKNSOGRLSPFKPATTSAPL 359
DB 300 EPDEEGYKELTCDRQFNEERKVNCAKILKRSQAQTOVRLDSFFKTLPTSTPATMAA--
QY 355 KKEISKTSKAANKKIKAGG 376
DB 358 KKK-AEAKKSNKKAKTS 377

RESULT 6
522267

hypothetical protein YK113C - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein, YK113C; RAD2 protein homolog

C:Species: Saccharomyces cerevisiae

C>Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 24-Sep-1995

C:Accession: S22267; S37941

R:Jacquier, A.; Legrain, P.; Dujon, B.

Yeast 9, 121-132, 1992

A>Title: Sequence of a 10.7 kb segment of yeast chromosome XI identifies th

5.

A:Reference number: S22267; M01D:92221689; PMID:1561835

A:Accession: S22267

A:Molecule type: DNA

A:Residues: 1-392 <YAC>

A:Cross-references: GB:93604; NID:9448391; P1DN:AA821998.1; P1D:9448392

A:Experimental source: strain S288C

R:Jacquier, A.; Legrain, P.; Gilleaux, L.; Richard, G.F.; Thierry, A.; Dujon

submitted to the protein sequence database, March 1994

A:Reference number: S37938

A:Accession: S37941

A:Molecule type: DNA

A:Residues: 1-382 <YAC>

A:Cross-references: EMBL:Z28113; NID:9448189; P1DN:CAA81953.1; P1D:94486190

A:Experimental source: strain S288C

C:Genetics:

A:Gene: SCD; RAD27

A:Cross-references: SGD:S0001596; MIPS:YK113C

A:Map position: 11L

C:Superfamily: conserved hypothetical protein YK113C

C:Keywords: nucleus

Query Match 48.4%; Score 938; DB 2; Length 382;
Best Local Similarity 49.5%; Pred. No. 20-55;
Matches 190; Conservative 69; Mismatches 109; Indels 16; Gaps

QY 1 MGKIGITKLADNAAKAKKEQFESEYFGKRIAVDASMSIYQFLIVYGRTGHTLTNEAGE
DB 1 MGKIGITKLADNAAKAKKEQFESEYFGKRIAVDASMSIYQFLIVYGRTGHTLTNEAGE
QY 61 VISHLQGMENRTIRLEAGIKPYVYEDGKPPDMKQOELAKRYSKRDATKDLTAVENGD
DB 61 TSHLMGMEYRTIRLDNDGKIPVYEDGKPPDLKSOELAKRERREAEKALKAAATDAD
QY 121 KDAIEKLSKRTVYKTRQHNEDCKRLRLMGVPEVPEAPSEAEACALCINDVFAVASED
DB 121 K--MKQERKLVVSKVSEHNEAQKLTGMDQIPITAPTEAEAGCALAKKGVYVAASED

181 MDLSTGAPRLHIMDPSSKIPVMEFYAKVLEFLTMDQFDLCTCGGDDYDJK 240
 Db 173 MDLSTGAPRLHIMDPSSKIPVMEFYAKVLEFLTMDQFDLCTCGGDDYDJK 237
 QY 241 GIGGOTALKLIRHOGSIESILEN-----NKKRYQIPEDMPYQEARLCEKPNWLTDFE 235
 Db 238 GVGPTATLKLIRHOGSIEKIVEFEESGSSNTKKIPEDMPYQAKMLDDEY-IGSNF 296
 QY 236 --LKWTPADEGLISLVKONGNEDRVTKAIEKTSANKSSQGRLESEFPYNT----- 350
 Db 297 INLKWSPKKEIELEYCDCKKSEERVKSGISRLKGLKSGIOGRDQGFVYVPRKIFQ 356
 QY 351 -SAPLKRRETSDDTKSAANKTKR 373
 Db 357 LAAAKRAQENKTKL-NKKNVTKR 379

RESULT 7 T46893

endo/exonuclease fen-1 (imported) - Pyrococcus furiosus
 N:Alternate names: DNA repair protein
 C:Species: Pyrococcus furiosus
 C:Date: 17-Mar-2000 *sequence_revision 17-Mar-2000 *text_change 21-Jul-2000
 C:Accession: T46893
 R:Disruptor: J. Brown, J. R. Bogert, A. P. Robb, F. T.
 J. Mol. Evol. 49, 474-484, 1999
 A>Title: DNA repair systems in Archaea: memories from the last universal common ancestor
 A:Reference number: 224126; MID:1994.5851; PMID:10486005
 A:Accession: T46893
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-340 <DIR>
 A:Cross-references: EMBL:AF013497; NID:q4102601; PID:AA01514.1; PID:q4102602
 A:Experimental source: strain DSM 3638
 C:Genetics:
 A:Gene: fen-1
 C:Superfamily: conserved hypothetical protein YKL113c

Query Match 33.4%; Score 648.5; Db 2; Length 340;
 Best Local Similarity 41.3%; Pred. No. 4, 1e-36;
 Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;

QY 19 KEQKFSYFGKRIADVAMS-YQPIVVGRTGMETLTNEGQVTSIHOGMNTTILFA 78
 Db 12 KEILENLYGKRIADVAMS-YQPIVVGRTGMETLTNEGQVTSIHOGMNTTILFA 71
 QY 79 GIKPYYVDGKPPDMKQKQETAKRYSKRQDAKDLTEAVEVGDKAIEKLSKRTVYKTRQ 138
 Db 72 GIKPYYVDGKPPDMKQKQETAKRYSKRQDAKDLTEAVEVGDKAIEKLSKRTVYKTRQ 131
 QY 139 NEDCKRLRLMGVYVVEAPSAEACALCTNKEVFAVEDMDSLTEGAPRCLRLMP 198
 Db 132 EDCAKLELMGIPYVAPSEGAAYMAKGSVASASODYS-JLFGPRIVRN-IT 193
 QY 199 SSKRIP-----VWEFPAKVALLEELTMQPIIDLCLICGQV-CDISKIQC 244
 Db 192 GKRLPGKNYVYKPELITL-----EVKLKELTKRKLIELMIVGTCVNPQGISIGL 247
 QY 245 QIALKLIRHOGSIESILENKKD--RYQIPEDMPYQEARLCEKPNWLTDFE 301
 Db 248 KRALELIVRH-----SKDPLAKFOKQSDVDLVAIKFEFLNPPV-ENYVLYVPR 295
 QY 302 DEGLISLVKONGNEDRVTKAIEKTSANKSSQGRLESEFPK 345
 Db 296 DEGLISLVKONGNEDRVTKAIEKTSANKSSQGRLESEFPK 339

RESULT 8 E75117

DNA repair protein rad2 (rad2) PAB1877 - Pyrococcus abyssi (strain Ozsasy)
 C:Species: Pyrococcus abyssi
 C:Date: 20-Aug-1999 *sequence_revision 20-Aug-1999 *text_change 20-Jun-2000

C:Accession: E75117
 R:anonymous; Genoscope
 Submitted to the EMBL Data Library, July 1999
 A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chr
 A:Reference number: A75001
 A:Accession: E75117
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-343 <KAW>
 A:Cross-references: GB:A124828; GB:A105836; NID:55459067; PID:CA649654.1
 A:Experimental source: strain Ozsasy
 C:Genetics:
 A:Gene: PAB1877
 C:Superfamily: conserved hypothetical protein YKL113c

Query Match 32.8%; Score 636; Db 2; Length 343;
 Best Local Similarity 41.6%; Pred. No. 2, 8e-35;
 Matches 143; Conservative 58; Mismatches 115; Indels 28; Gaps

QY 19 KEQKFSYFGKRIADVAMS-YQPIVVGRTGMETLTNEGQVTSIHOGMNTTILFA 78
 Db 12 KEILENLYGKRIADVAMS-YQPIVVGRTGMETLTNEGQVTSIHOGMNTTILFA 71
 QY 79 GIKPYYVDGKPPDMKQKQETAKRYSKRQDAKDLTEAVEVGDKAIEKLSKRTVYKTRQ 138
 Db 72 GIKPYYVDGKPPDMKQKQETAKRYSKRQDAKDLTEAVEVGDKAIEKLSKRTVYKTRQ 131
 QY 139 NEDCKRLRLMGVYVVEAPSAEACALCTNKEVFAVEDMDSLTEGAPRCLRLMP 198
 Db 132 EDCAKLELMGIPYVAPSEGAAYMAKGSVASASODYS-JLFGPRIVRN-IT 193
 QY 199 SSKRIPMEVY-----AKLELEELTMQPIIDLCLICGQV-CDISKIQC 244
 Db 192 GKRLPGKNYVYKPELITL-----EVKLKELTKRKLIELMIVGTCVNPQGISIGL 247
 QY 245 QIALKLIRHOGSIESILENKKD--RYQIPEDMPYQEARLCEKPNWLTDFE 301
 Db 248 KRALELIVRH-----SKDPLAKFOKQSDVDLVAIKFEFLNPPV-ENYVLYVPR 295
 QY 302 DEGLISLVKONGNEDRVTKAIEKTSANKSSQGRLESEFPK 345
 Db 296 DEGLISLVKONGNEDRVTKAIEKTSANKSSQGRLESEFPK 339

Query Match 32.6%; Score 632.9; Db 2; Length 343;
 Best Local Similarity 39.2%; Pred. No. 4, 8e-35;
 Matches 142; Conservative 66; Mismatches 110; Indels 35; Gaps

QY 1 MCIGKLTLLADNAKMKQKQETAKRYSKRQDAKDLTEAVEVGDKAIEKLSKRTVYKTRQ 138
 Db 1 MSVP-----IDDLVPR-KEIDLEMLYGKRIADVAMS-YQPIVVGRTGMETLTNEGQVTSIHOGMNTTILFA 78
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic archaeon, Pyrococcus horikoshii
 A:Reference number: A71009; MID:98344137; PMID:9579194
 A:Accession: A71015
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-343 <KAW>
 A:Cross-references: GB:AF000006; NID:q4126133; PID:BA03621.1; PID:q4126133
 A:Experimental source: strain 013
 A:Note: This accession replaces an earlier accession for a sequence replaced
 C:Genetics:
 A:Gene: PH145
 C:Superfamily: conserved hypothetical protein YKL113c

```

OY 61 VTSHLQGMENRTLRLEAGIKPYVVDGKPPDMKQOELAKRYSKRDNDATKDLTEAVGSD 120
DB 54 ITSHLSGLTYRTINLMEAGIKPAIVFDGKPPDKRELEKREAREAEALKKKEALAKGN 113
OY 121 KDAIEKLSKRTVYTRQHNEDECKRLRLMGVPEVPEASEAEACALCINDVFAVASPD 180
DB 114 LEFARKYAKRATKYNMLLEDAKMLLQMLGPIILQAPSGEQAQVAMSKGVYASADQ 173
OY 181 MDLTFGARPRFLRLHMDPSSKKIPVMEFDVA-----KYLELELTMDQFDLCILC 231
DB 174 YSLLEFGARPLRLNLTITCKRMPKPGKDVYEIKPELVADVFLKELIKITREKLELAIIV 233
OY 232 GCDY-CDS:KGIGGQATLKLIRHGSISLELNKD---RYQIEDMPYQARSLKEP 267
DB 234 GTDYNPGVGKIGGPKALEIVR-----YSNDPLAKPQSDVDVLAIKKEFLNP 282
OY 288 NVTLDPLELWPAPEEGELISFLVKNNGENEDRVTKAIKRSKAKKSSQGRLESF--- 344
DB 283 PVINRY-SLSKKEPPDEGLIKFLCDEPHNSERVRKNGIERLKKALKAGQSTLPSRWYKK 341
OY 345 KP 346
DB 342 KP 343

```

RESULT 10

```

CG64480
DNA repair protein RAD2 homolog - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C:Accession: C64480
R:Belt, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischman, R.D.; Sutton, G.G.; Blake,
R.; Helic, C.L.; Overbeek, R.; Kikness, E.F.; Weissstock, K.G.; Kierick, J.K.; Goodex, A.;
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.;
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Kienk, H.P.; Fraser, C.M.; Smith, H.O.; Moese,
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A:Reference number: A64300; MIMD:56337993; PMID:8688937
A:Accession: C64480
A:Status: Preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-326 <KLE>
A:Cross-references: GB:067585; GB:077117; MIMD:562088; PIDN:AB39454.1; PID:592090;
C:Genetics:
A:Map position: FCRI41334-14.4314
C:Superfamily: conserved hypothetical protein YKL113c

```

```

Query Match 28.9%; Score 560.5; DB 2; Length 326;
Best local similarity 38.0%; Pred. No. 36-30;
Matches 132; Conservative 71; Mismatches 121; Indels 23; Gaps 9;

OY 1 MGKIGLITKLADNAPKAMKQKFESEYFGKKIADVADSMSTYQFLIYVGRGTMLNEAGE 60
DB 1 MGVO-----FQDFPKNI--ISFDIKGKKVAIDGNALYQFLISRLRQSGPLNKRGE 53
OY 61 VTSHLQGMENRTLRLEAGIKPYVVDGKPPDMKQOELAKRYSKRDNDATKDLTEAVGSD 120
DB 54 ITSAVGVYFYKTIHLEND:TLVWEDGEPFLKRETRVRRMKKAKKKKEIKKMD 113
OY 121 KDAIEKLSKRTVYTRQHNEDECKRLRLMGVPEVPEASEAEACALCINDVFAVASPD 180
DB 114 LEFARKYAKRATKYNMLLEDAKMLLQMLGPIILQAPSGEQAQVAMSKGVYASADQ 173
OY 181 MDLTFGARPRFLRLHMDPSSKKIPVMEFDVA-----KYLELELTMDQFDLCILC 231
DB 174 YSLLEFGARPLRLNLTITCKRMPKPGKDVYEIKPELVADVFLKELIKITREKLELAIIV 233
OY 232 GCDY-CDS:KGIGGQATLKLIRHGSISLELNKD---RYQIEDMPYQARSLKEP 267
DB 234 GTDYNPGVGKIGGPKALEIVR-----YSNDPLAKPQSDVDVLAIKKEFLNP 282
OY 288 NVTLDPLELWPAPEEGELISFLVKNNGENEDRVTKAIKRSKAKKSSQGRLESF--- 344
DB 283 PVINRY-SLSKKEPPDEGLIKFLCDEPHNSERVRKNGIERLKKALKAGQSTLPSRWYKK 341
OY 345 KP 346
DB 342 KP 343

```

```

DB 280 LPDEGILKPLVDEPNDRYDSYKKHYAKITNLINKTKQKTLQWEX 326

```

RESULT 11

```

CG6285
DNA repair protein Rad2 - Methanobacterium thermoautotrophicum (strain Del.)
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 24-Sep-1999
C:Accession: C69085
R:Smith, D.R.; Deucette-Stamm, L.A.; Delonghery, C.; Lee, H.; Dubois, C.;
Oiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Mierzbowski, J.; Gibson, J.;
K.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Neelling, J.; Reedy
Bacteriol. 179, 7185-7195, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum
A:Reference number: A69000; MIMD:98037514; PMID:9371463
A:Accession: C69085
A:Status: Preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-328 <KMT>
A:Cross-references: GB:AE00922; GB:AE00666; MIMD:92622754; PIDN:AA86106.
A:Experimental source: strain Delta H
A:Genetics:
A:Gene: MIM1633
C:Superfamily: conserved hypothetical protein YKL113c

```

```

Query Match 28.0%; Score 542; DB 2; Length 328;
Best local similarity 36.5%; Pred. No. 528-29;
Matches 126; Conservative 73; Mismatches 148; Indels 18; Gaps

```

```

OY 1 MGKIGLITKLADNAPKAMKQKFESEYFGKKIADVADSMSTYQFLIYVGRGTMLNEAG 120
DB 1 MGKV-----LKVVSFPI---RLEDLGRKVAADANTTIQFSSIFQSGTLMNSRG 113
OY 61 VTSHLQGMENRTLRLEAGIKPYVVDGKPPDMKQOELAKRYSKRDNDATKDLTEAVGSD 180
DB 53 EYTSHLQGMENRTLRLEAGIKPYVVDGKPPDMKQOELAKRYSKRDNDATKDLTEAVGSD 173
OY 121 KDAIEKLSKRTVYTRQHNEDECKRLRLMGVPEVPEASEAEACALCINDVFAVASE 231
DB 113 KTSLSHLYIRTAAMERIRVYFDGPRSHMLKCEVSEKADLRKSEVEMKRALEEG 223
OY 181 MDLTFGARPRFLRLHMDPSSKKIPVMEFDVA-----KYLELELTMDQFDLCILC 231
DB 173 EYTSHLQGMENRTLRLEAGIKPYVVDGKPPDMKQOELAKRYSKRDNDATKDLTEAVGSD 223
OY 232 GCDY-CDS:KGIGGQATLKLIRHGSISLELNKD---RYQIEDMPYQARSLKEP 267
DB 212 KQIQAARMKIKLREKQIEKVLKQVADIGDP-----QVLRPRFLEPVESEY-ELRWR 256
OY 288 NVTLDPLELWPAPEEGELISFLVKNNGENEDRVTKAIKRSKAKKSSQGRLESF 344
DB 286 KPIVEGVLEFLTEHGFSEQVYFAALKKEFA--STVKSLEDMF 326

```

RESULT 12

```

CG69282
DNA repair protein Rad2 (rad2) homolog - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 24-Sep-1999
C:Accession: H69282
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Kotchum
A: Fleischman, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.;
A:Authors: Overbeek, R.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reduc
A:Reference number: A69250; MIMD:98049143; PMID:9389475
A:Accession: H69282
A:Status: Preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-336 <KLE>

```

A:Cross-references: G3:AE001087; G8:AE000782; NID:q2588410; FIDN:BA090367.1; PID:q258537
 C:Superfamily: conserved hypothetical protein YK1130c

Query Match

27.8%; Score 538.5; DB 2; Length 336;

Best Local Similarity 36.7%; Pred. No. 9.2e-29;

Matches 125; Conservative 67; Mismatches 118; Indels 31; Gaps 7;

QY

19 KEQKESYFGKRIADVADMSITQFLIYVGRGKMTLTNAGEVTSILQGHNRITELIA 78

DB

12 EEVELETFSGKRIADADANTLYOFISITRQDPGLDSDGRIISHLISGLIYRSMVAV 71

QY

79 GIKPYVVEGKPPPMKKKQELAKRYKRDATKDTLEAVEGDDKAIKSLKRIKTVRCH 138

DB

72 GIRPVFVFGDPEPKKKEIEERKKRRALAEEMIALQGDNDK-KVYQAGRCVCEV 130

QY

139 NEDCKRLRLMGVVPVEAFSEAEACALCINDKVPVASEDMDSLTFFGAPRLRLMDP 198

DB

131 VDSAKTLLSYMGIPVDAPEGEQAAVMAKQCVETYSQDYDSLTFGSPRCARNLAI 190

QY

199 SSKKIP-----VMEFDVAKYLELELTINQFIDLCICGCEYCDG-KGIGGQ 245

DB

191 GKRKLPGKNVYVPKPIITILIESN-----LKRIGLTRQLDIALIVGTEVNGKRYGVK 246

QY

246 TALKLRQHGSIESTLENL--NKDRYQIPEDWYQARRLKEPNVTLDIPEKMTVAPE 303

DB

247 KALNYIKTYGDIFFALKALKVNIH-----VEEIRNFECNPPTIDCY-KIEFFREDF 297

QY

304 EGIISFLVKNDFNEDRYTKAIEKSKAKNSQGLIESFF 344

DB

298 EKAIEFLCEHDFSEKREKALEKLA--KSTQATLRRMF 336

RESULT 13

H72765

Probable flap endonuclease-1 APE0115 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C:Date: 20-Aug-1999 *sequence_revision 20-Aug-1999 *text_change 20-Jun-2000

C:Accession: H72765

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamaizaki, S.; Hatahara, Y.; Jin-no, K.; Takai-awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamaizaki, J.; K

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyperthermophilic Crenarchaeon, *Aeropyrum*

A:Reference number: A72450; MIMD:99310339; PMID:10382966

A:Accession: H72765

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-401 (KAN)

A:Cross-references: DBJ:AF000058; NID:q5103388; PIDN:BA079026.1; FID:q5103555

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE0115

C:Superfamily: conserved hypothetical protein YK1130c

Query Match

26.5%; Score 514.5; DB 2; Length 401;

Best Local Similarity 34.6%; Pred. No. 4.6e-27;

Matches 119; Conservative 74; Mismatches 122; Indels 19; Gaps 6;

QY

15 EKAKKQKPESTYFGRKIVADMSITQFLIYVGRGKMTLTNAGEVTSILQGHNRITELIA 74

DB

62 PEAREVELRALSGVYALDAYNMLOYPIITAIROPDGPILIDREGHVTSILSHGLEYPTN 121

QY

75 LLEAGIKPVVEGKPPPMKKKQELAKRYKRDATKDTLEAVEGDDKAIKSLKRIKTVRCH 134

DB

122 IYBERIKPVYFVFGDPEPKKKEIEERKKRRALAEEMIALQGDNDK-KVYQAGRCVCEV 130

QY

135 TQNNEDCKRLRLMGVVPVEAFSEAEACALCINDKVPVASEDMDSLTFFGAPRLRLMDP 198

DB

102 TSDVFESEKELIDANGKPMVQAPAEAGEQAAVMAKQCVETYSQDYDSLTFGSPRCARNLAI 190

QY

195 LWDPSKRIPIVVE-----FCVAKYLELELTINQFIDLCICGCEYCDG-KGIGGQ 244

DB

242 LATGRRKLPGRDQVVEIKETITELDELISGLTREGDLAVGILACTGYNIGGVGRVGP 301

QY 245 OTALLILQHGSIESTLENLKDKRYQIPEDWY-QEARRLKEPNVTLDIPEKMTVAPE 303

DB

302 KALNRYVSLGDPKPKVIAVPRGCV---DDVYLRKVEYELNPPVDDY-KIEFFREDF 297

QY

304 KOLISFLVKNDFNEDRYTKAIEKSKAKNSQGLIESFF 344

DB

357 DKVKEIVERRHDFNPERERALKERLAKRYKRLDGRSRLDMF 400

RESULT 14

F90158

DNA repair endo/exonuclease FEN-1 (RAD2) (RAD2) [imported] - Sulfolobus sol

C:Species: Sulfolobus solfataricus

C:Date: 24-May-2001 *sequence_revision 24-May-2001 *text_change 15-Jun-2001

C:Accession: F90158

R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, V.; Alward, G.; Aways

Jong, T.; Joffiles, A.C.; Kozera, C.J.; Medina, N.; Feng, X.; Thi-Ngoc, H.

arrett, R.A.; Ragan, M.A.; Seisen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A89119

A:Accession: F90158

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-302 (KOR)

A:Cross-references: G3:AE006642; NID:q381333; PIDN:AAK40525.1; GSPDB:GN00

C:Genetics:

A:Gene: rad2

C:Superfamily: conserved hypothetical protein YK1130c

Query Match

26.1%; Score 506; DB 2; Length 302;

Best Local Similarity 37.9%; Pred. No. 1.2e-26;

Matches 118; Conservative 58; Mismatches 97; Indels 38; Gaps

QY

59 GEVTHLQGMFRITRILLIAGIKPVYVDSKPPPMKKKQELAKRYKRDATKDTLEAVE 303

DB

5 GRVTHLSGLFRTINILEGVPIVYVDSKPPPMKKKQELAKRYKRDATKDTLEAVE 303

QY

119 GCKDAIHKISKRYVTVQHNEDCKRLRLMGVVPVEAFSEAEACALCINDKVPVAVAS 198

DB

65 GRIEELKRYSOAILRLSNINVEESKRLRAMSGPIVQARNSGFAFVYLNKLGSLMAAS 190

QY

175 EMDSLTGGARFELRLMDPSKIP-----VMEVAKVLELELTINQFID 190

DB

125 QDYDALTGAKRLVNIITGRKRLPNKDYVEIKPELITEL--LKKLITREOITDI 190

QY

228 GIIGCCDY-CDSIKGICGTALIKILQHGSIESTLE-----NKKDRYQIPEDWY 303

DB

183 GILISTGNPDIIRIGTSPFRALIKIKKQIKKAKVEYELISKDINRILD----- 303

QY

278 QAKKFKPEPNVTLDIPIKMTAHEHGTIFELVKNDFNEDRYTKAIEKSKAKNSQGLIESFF 344

DB

233 ELRGLFNPVQVKEFEALDNPNGEDLNLVYEHNFSEKRVKNGIERLTKAIKAKG 344

QY

334 KSGQGLIESFF 344

DB

292 AEGQNLDMF 302

RESULT 15

F84290

DNA repair protein (imported) - Halobacterium sp. NRC-1

C:Species: Halobacterium sp. NRC-1

C:Date: 02-Feb-2001 *sequence_revision 02-Feb-2001 *text_change 16-Feb-2001

C:Accession: F84290

R:Katz, M.V.; Kennedy, S.P.; Matsuzaki, S.G.; Bergquist, B.; Fan, M.; Shukla, H.

Leibman, B.; Keller, K.; Cruz, R.; Danesh, M.J.; Hough, D.W.; Maddocks

Jung, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Katz, S.P.; Danesh, M.J.; Dennis, P.P.; Omer, A.D.; Fbhardt, H.; LC

A:Title: Genome sequence of Halobacterium species NRC-1

A:Reference number: A81160; MIMD:20504483; PMID:11016950

A:Accession: F84290

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-327 <STO>
 A:Cross-references: CB:AE004437; NID:q13580870; PIDN:AA319690.1; GSP3:GN00138
 C:Genetics:
 A:Gene: rad2
 C:Superfamily: conserved hypothetical protein YKL113c

Query Match 19.0%; Score 369; D5 2; Length 327;
 Best Local Similarity 31.8%; Pred. No. 1.9e-17;

Matches 106; Conservative 50; Mismatches 139; Indels 28; Gaps 9;

```

QY 17 AKKEQKESYFGRIADVASMTYQFL-IYVGRTGNETLTNEAGEVTSHTQGFNRTIRL 75
   |::| | |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db 11 ALEETPFADLEGSVAADAHNMILKYLTTVQWAGADVTTSDQTEVANLWCAVAGCLPKR 70
   |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
QY 76 LEAGIKPYVPFDGKPPDKKQGLAKRYSKRDDATKDLTFAVEVGDKDALFKLSKRTVXYT 135
   |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db 71 FEHGLTPVFWMDGVTETIKDEIADREQREYEEQLDAREAGDAEAARLDAFTQRLT 130
   |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
QY 136 RQHNEQCKRLRLKMGVYVEAPSEAEACALC-INUKVFAVASEDMDSLTPGATFLER 194
   |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db 131 PTIHETRELELDLDPQVEAPAEGEAAVMTRIDPAVDYAGSDDYDCLLSPTLTNQ 190
   |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
QY 195 LMDPSKRIPVMEFPAKVLFELEFIMQFICILQCDYCDYSIKGIGQFALKLIRQH 254
   |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db 191 L--TSSCHPEIMFDA--TCAEHDLTWEQLVDVGLCGTFNPGIDGFCPTALDAGCH 246
   |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
QY 255 GSIESIL---ENL-NKDRYQIPEDWPIQFARRLFRKPNVTLDI---PELKNATATDEEG 305
   |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db 247 GDLMCVLAEGEHHVHGDR-----LRELFINPDVTDVYIDGV---SFAIDAA 292
   |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
QY 307 ISFLVKNQGFNEEDRYTKAIEKIKSAKNSQGR 339
   |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db 293 RAVVTDEMEVDADAVARGFERIDMAAAGTGLNR 325
   |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|

```

Search completed: October 22, 2003, 12:43:25
 Job time : 44 secs

LR SMART: SMOU435; XPGN: 1.
 DR PROSITE: PCS0841; XPG_1: 1.
 CR PROSITE: PCS0842; XPG_2: 1.
 KW Hydrolase; Nuclease; Endonuclease; Nuclear protein.
 FT DOMAIN 1 102 N-DOMAIN.
 FT DOMAIN 120 251 I-DOMAIN.
 SU SEQUENCE 376 AA: 42314 MW: 827946HDBDFE9 CRG04;

Query Match 52.3%; Score 10.4; DR 1; Length 376;
 Best Local Similarity 53.2%; Pred. No. 2.9e-69;
 Matches 202; Conservative 69; Mismatches 101; Indels 8; Gaps 6

QY 1 MGKIGKTRLLANAMKAMKEQKPFSEYGRKAVNASISYQILIVGRIGMENTNAGE 50
 1 MGKIGKTRLLANAMKAMKEQKPFSEYGRKAVNASISYQILIVGRIGMENTNAGE 50
 Db 1 MGKIGKTRLLANAMKAMKEQKPFSEYGRKAVNASISYQILIVGRIGMENTNAGE 59
 1 MGKIGKTRLLANAMKAMKEQKPFSEYGRKAVNASISYQILIVGRIGMENTNAGE 59
 QY 61 VTSLQGMENRTIRLEAGIKPVYVFDGKPPDMKKOELARYSKRDPATKDLTAEVGD 120
 61 VTSLQGMENRTIRLEAGIKPVYVFDGKPPDMKKOELARYSKRDPATKDLTAEVGD 120
 Db 60 TTS-LMGMEYRTIR-MENCIKPVYVFDGKPPDMKKOELARYSKRDPATKDLTAEVGD 117
 60 TTS-LMGMEYRTIR-MENCIKPVYVFDGKPPDMKKOELARYSKRDPATKDLTAEVGD 117
 QY 121 KQAIERLSRRTKVTROHNEQCKRLRLRMGYPVVAESEAFAEALCIDKRYFAVASED 180
 121 KQAIERLSRRTKVTROHNEQCKRLRLRMGYPVVAESEAFAEALCIDKRYFAVASED 180
 Db 118 EEEVEKTRRLYKVKQHNDECKHLLSMGIPYLDAPSEAFASCAALAKAKYAAATED 177
 118 EEEVEKTRRLYKVKQHNDECKHLLSMGIPYLDAPSEAFASCAALAKAKYAAATED 177
 QY 181 MDSTFGARFLRLHMDPSSKIPMEEDYAKVLEELMTDQFDLCLAGGDCYSIR 240
 181 MDSTFGARFLRLHMDPSSKIPMEEDYAKVLEELMTDQFDLCLAGGDCYSIR 240
 Db 178 MGLTFGSPVLMRHLTLASEAKKLPIQEFHLSRYLQELGLNQEOFDVCLILGSDYCSIR 237
 178 MGLTFGSPVLMRHLTLASEAKKLPIQEFHLSRYLQELGLNQEOFDVCLILGSDYCSIR 237
 QY 241 GIGGQATALKIQHQSIESILENLKRDYQIPEDMPYQEARLFFKEPNNT-LDIPELKMT 299
 241 GIGGQATALKIQHQSIESILENLKRDYQIPEDMPYQEARLFFKEPNNT-LDIPELKMT 299
 Db 238 GIGARAVDILQKHSIDELRYRLDPKSYPPPEWMLKREKQQLFLEEYVDPSPVELKKS 297
 238 GIGARAVDILQKHSIDELRYRLDPKSYPPPEWMLKREKQQLFLEEYVDPSPVELKKS 297
 QY 300 APDEGLISFLYKONGFNEDRYTAIEKINSKAKSSQGRLESEFKPTATAPLARKET 359
 300 APDEGLISFLYKONGFNEDRYTAIEKINSKAKSSQGRLESEFKPTATAPLARKET 359
 Db 298 ENEEELVYFMGEGKQFSEERIRSGVKRLSKSRGSGYGRILDPEFKYGLSL-ARKKEP 356
 298 ENEEELVYFMGEGKQFSEERIRSGVKRLSKSRGSGYGRILDPEFKYGLSL-ARKKEP 356
 QY 360 SDKTSAAANKTKKAGKKK 379
 360 SDKTSAAANKTKKAGKKK 379
 Db 357 E---PKGPARRKAKTGAGK 373
 357 E---PKGPARRKAKTGAGK 373

RESULT 4
 RA27_YEAST
 ID RA27_YEAST STANDARD: PRT: 382 AA.
 AC 326753;
 DT 0: AUG-1992 (Rel. 23, Created)
 DT 0: AUG-1992 (Rel. 23, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Structure specific endonuclease RAD27.
 GN RAD27 OR RTH1 OR YKL113C OR YKC510.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.
 CC NCBI:taxid:4932;
 RN 11
 RP SEQUENCE FROM N.A.
 RX MEDLINE:92221689; PubMed:1561835;
 RA Jacquier A., Legrain P., Dujon B.;
 RT "Sequence of a 10.7 kb segment of yeast chromosome XI identifies the
 ABNI and the Baf1 loci and reveals one tRNA gene and several new open
 PT reading frames including homologs to RAD2 and kinases."
 RL Yeast 8:121-132(1992).
 RN 12
 RP CHARACTERIZATION.
 RC STRAIN:SAHBC;
 RX MEDLINE:95011546; PubMed:7926735;
 RA Harrington J.J., Lieber M.R.;
 RT "Functional domains within FEN-1 and RAD2 define a family of
 KI structure-specific endonucleases: implications for nucleotide
 RC excision repair."
 RL Genes Dev. 8:1344-1355(1994).
 RN 13
 RP CHARACTERIZATION.

```

RX MEDLINE:9501133773; PubMed 7834425.
RA Haagen M.S., Pittenger C., Slade W., Friedberg S.C.:
R1 "Characterization of a mutant strain of Saccharomyces cerevisiae with
R2 a deletion of the Rad27 gene, a structural homolog of the Rad2
R3 nucleotide excision repair gene."
R4 J. Bacteriol. 177:364-371(1995).
CC -1- FUNCTION: SINGLE-STRANDED DNA ENDONUCLEASE (PROBABLE).
CC -1- SUBCELLULAR LOCATION: Nucleol (Probable).
CC -1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1
CC SUBFAMILY.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaborat
CC between the Swiss Institute of Bioinformatics and the EMBL outsta
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no
CC modified and this statement is not removed. Usage by and for comm
CC entities requires a license agreement (See http://www.isb-sib.ch/annou
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: S93804; AAC31998.1; .
DR EMBL: Z28113; CAAB1953.1; .
DR PIR: S22677; S22677.
DR HISSP: Q58839; IAT6.
DR GK: P26793; .
DR SGD: S0001596; RAD27.
DR GO: GO:0005634; C:nucleus; IMP.
DR GO: GO:0017-08; F:flap endonuclease activity; IDA.
DR GO: GO:0000731; P:DNA repair synthesis; IMP.
DR InterPro: IPR002421; 5_3_exonuclease.
DR InterPro: IPR000513; Exo_N_1.
DR InterPro: IPR003584; HNH_2.
DR InterPro: IPR006086; XPG_1.
DR InterPro: IPR006085; XPG_N.
DR InterPro: IPR006084; XPGC_Rad.
DR Pfam: PF00867; XPG_1; 1.
DR Pfam: PF00752; XPG_N; 1.
DR PRINTS: PR00683; XPGRADSUPER.
DR SMART: SM00279; HHE2; 1.
DR SMART: SM00484; XPG1; 1.
DR SMART: SM00485; XPGN; 1.
DR PROSITE: PS00841; XPG_1; 1.
DR PROSITE: PS00842; XPG_2; 1.
KW DNA repair; Hydrolase; Nuclease; endonuclease; Nuclear protein.
FT DOMAIN 1 105 N-DOMAIN.
FT DOMAIN 120 251 I-DOMAIN.
FT SREGIONCE 382 AA; 43279 MM; I54B097202129C CRO64;
CC -----
Cherry Match 48.4%; Score 948; DE 1; Length 382;
Host Local Similarity 49.5%; Prod. No. 3.2e-55;
Matches 190; Conservative 69; Mismatches 139; Indels 16; Gaps
CC -----
QY 1 MGKIGITLLALNANRAKAKKEDKPESTYQKRIAVYASMS;IGELIVVGRTSMKELTINAGR 6
Db 1 MGKIGITLLALNANRAKAKKEDKPESTYQKRIAVYASMS;IGELIVVGRTSMKELTINAGR 6
QY 6: VSHLDGMRNRIIRLEAGTIRKVFVEFCJRPDMKQELARYSKREDADK;TEVEVWD 1
Db 6: VSHLDGMRNRIIRLEAGTIRKVFVEFCJRPDMKQELARYSKREDADK;TEVEVWD 1
QY 61 TISHLMGMRYPLRLN;DNGIKTCYVPGPKRPDLKSHITFRSSRYRETEKLAALATLE 1
Db 61 TISHLMGMRYPLRLN;DNGIKTCYVPGPKRPDLKSHITFRSSRYRETEKLAALATLE 1
QY 121 KDA;EKLKSRIVYK;VROHNEDCKRI;R;MGVYVPEAPSPAPDCAALCINIVYFVASHD 1
Db 121 KDA;EKLKSRIVYK;VROHNEDCKRI;R;MGVYVPEAPSPAPDCAALCINIVYFVASHD 1
QY 121 K--MKQSRRLKLVKSKRENEADKIKLS;MGIPY;IATPFAEACAVELAKRGVYAAASE 1
Db 121 K--MKQSRRLKLVKSKRENEADKIKLS;MGIPY;IATPFAEACAVELAKRGVYAAASE 1
QY 161 MDSITGCAIFELRYL;MDSYSKRIYVNS;FYAKYV;FELETLMDGIFDCLTCGDCGDSIK 2
Db 161 MDSITGCAIFELRYL;MDSYSKRIYVNS;FYAKYV;FELETLMDGIFDCLTCGDCGDSIK 2
QY 176 MTLTCGRTFFILAKHLTSEAKKEP;IHELDTFLVINKGLD;IIPQVYD;C;MLGDCGDSIR 2
Db 176 MTLTCGRTFFILAKHLTSEAKKEP;IHELDTFLVINKGLD;IIPQVYD;C;MLGDCGDSIR 2
QY 241 G;GGGATLALKLRQGSIESILEN;----NKDRIYV;PEMPVYEARLCKEPVYIDJPE 2
Db 241 G;GGGATLALKLRQGSIESILEN;----NKDRIYV;PEMPVYEARLCKEPVYIDJPE 2
QY 238 GGGPVPYALK;CTITSS;TKIVLFFIFSGSNSNTKK;ITEDIPYVQAQMLLDPEV;LDGNE 2
Db 238 GGGPVPYALK;CTITSS;TKIVLFFIFSGSNSNTKK;ITEDIPYVQAQMLLDPEV;LDGNE 2
QY 296 --LKTATDEGRLISFLVKGDPEDRYKATKIRISAKKSSQGLSEFPKPTATT---
Db 296 --LKTATDEGRLISFLVKGDPEDRYKATKIRISAKKSSQGLSEFPKPTATT---

```

```

Dn 297 : NIKWSPKRELLIEYLCDOKKFSFEERKKSISRLKGLKSGIQGRLDGFFVYVPKTAEQ 356
Qy 351 -SAPLKRRETSOKTSKAANKTK 373
Db 357 LAAAKRAOENKKNLKN -NKNKVTK 379

RESULT: 5
FEN_PyFU
ID_FEN_PyFU STANDARD: PRT: 340 AA.
AC 093634;
DI 28-FEB-2003 (Rel. 41, Created)
DI 28-FEB-2003 (Rel. 41, Last annotation update)
DI 28-FEB-2003 (Rel. 41, Last annotation update)
DE Flap structure-specific endonuclease (EC 3.1.1.1).
GN FEN OR FEN1 OR Pfl14.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
Pyrococcus.
NCBI_TaxID=2261;
..
RP SEQUENCE FROM N.A.
RC STRAIN-VCL / DSM 3638 / ATCC 43587 / JCM 8422;
RX MEDLINE-99415851; PubMed-10486005;
RA Dlugosz J., Brown J.R., Robert A.P., Robs F.T.:
RC "DNA repair systems in archaea: mementos from the last universal
RT common ancestor?";
RL J. Mol. Evol. 49:474-484(1999).
..
RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX HOSFIELD D.J., MOL C.D., SHEN B., TAIHER J.A.;
MEDLINE-98449703; PubMed-9778254;
RT "Structure of the DNA repair and replication endonuclease and
RL exonuclease FEN1: coupling DNA and PCNA binding to FEN1 activity.";
Cell 95:115-146(1998).
..
CC -1. FUNCTION: Endonuclease that cleave the 5'overhanging flap
CC structure that is generated by displacement synthesis when DNA
CC polymerase encounters the 5'end of a downstream Okazaki fragment.
CC Has 5'endo-/exonuclease and 5'pseudo-Y-endonuclease activities.
CC Cleaves the junction between single and double-stranded regions of
CC flap DNA.
CC -1. COFACTOR: Binds 2 magnesium ions per subunit.
CC -1. SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1
CC SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/sib-stb.ch).
CC or send an email to license@sib-sib.ch).
CC
DR EMBL: AF013497; AAC01514.1; -
DR EMBL: AF010244; AA81538.1; -
DR PIR: T46893; T46893.
DR PDB: 1B43; 12-JAN-00.
DR HAMAP: MF_00614; -; 1.
DR InterPro: IPR000513; EXO_N_1.
DR InterPro: IPR003584; HHR_2.
DR InterPro: IPR006086; XPG_1.
DR InterPro: IPR006085; XPG_N.
DR InterPro: IPR006084; XPGC_Rad.
DR Pfam: PF00867; XPG_T; 1.
DR Pfam: PF00752; XPG_N_1.
DR PRINTS: P80083; XPGRADSUPPER.
DR SMART: SMC0279; Hhh2; 1.

```

```

DR SMART: SMO03484; XPG1: 1.  

DR SMART: SMO03484; XPG1: 1. FALSE_NEG.  

DR PROSITE: PS006841; XPG_1:  

KM Hydroxylase: Nuclease; Endonuclease; Magnesium; Metal-binding;  

RM 3D-structure; Complete proteome.  

KT METAL 154 154 MAINESEUM I (BY SIMILARITY).  

SQ SEQUENCE 340 AA: 38738 NM: D8H3D0F999550P CPE64:  

  

Query Match 33.48; Score 648.5; DB 1: Length 340;  

Best Local Similarity 41.38; Pred. No. 4,3e+36;  

Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps  

  

QY 19 KKKFFSEYERKAVADSKSYIOELLVSGRTGETLTNENGEVTSHLOGMFNRIITLLEA  

DB 12 KEIELENLYRKAIATIDANAYIFLSTIQOKOTPLAMDCKRITSLSOLFPTIINLMRA  

QY 79 GKRVVVEDGRPPDMKRKQELARKYSRKDATIKDTSAHVGSMDA;FKISKRPKVVSQH  

DB 72 GKIKVVVFVDDEPPFPFKLEKEREREAEAEEREALEKELEEARKNVAQATVNVNL  

QY 139 NEGKKRLIRLMGVVVVHAHSFAEAGNALCTINKVPAVANECDKSLTHAPRFLLHMDP  

DB 132 TEQDKKLLELMGIPVCAHSNGCALAAVMAMAKSVASASOYDSLTPAPSPVRLNTTT  

QY 199 SSRRKP-----VMPEDVAKVFELETITDQCETDICTICQGDY-CDSIKSQQ  

DB 192 GKRLPRKNVVELKPELLIIH---HWKELKLTNEKILELATLVGTGYNRGGIKGLDL  

QY 245 QTALKLIKLGHSLESTLENLNKC---RYCPEHPVGLARLPKRENVTLDPKLKWIAE  

DB 248 KKALEIVR-----SKLELRFSKSDMTVAIKREFLINPVTDVYNLVMWCP  

QY 302 DEETHLEPKKCNQPNFERVTKAFEXFKSAKSSSGRLSEPFK 345  

DB 296 DEBITKETCEFDNPFRFKKMLHLKKA.NSKSKSTLESMEK 339  

  

RESULTS  

FEN_FEN_PYRAH STANDARD; PRT: 343 AA.  

ID FEN_PYRAH  

AC Q9Y0Z9;  

CF 28-FEB-2003 (Ref. 41, Created)  

DF 28-FEB-2003 (Ref. 41, Last sequence update)  

DT 15-SEP-2003 (Ref. 42, Last annotation update)  

DE Flap structure-specific endonuclease (EC 3.-.-.-).  

GN FEN OR PYRAD074C0 OR PAH1477.  

OS Pyrococcus abyssi.  

OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  

CC Pyrococcus.  

CX NCBI_TaxID=29292;  

NM RP  

RP SEQUENCE FROM N.A.  

RC STRAIN-GHS / Uruguay;  

RX PubMed=12622808;  

RA Cohen G.N., Barbe V., Flament D., Galerio M., Heilig R., Lecompte O.,  

RA Poeh O., Prieur D., Querellou J., Zivanovic Y., Forterre P.,  

RA Van der Post J., Weissenbach J., Zivanovic Y., Forterre P.,  

RT "An integrated analysis of the genome of the hyperthermophilic  

RT archaeon Pyrococcus abyssi." ;  

RL Mol. Microbiol. 47:1495-1512(2003).  

CC -FUNCTION: Endonuclease that cleave the 5'overhanging flap  

CC structure that is generated by displacement synthesis when DNA  

CC polymerase encounters the 5'end of a downstream Okazaki fragment.  

CC Has 5'endo-/exonuclease and 5'pseudo-Y-endonuclease activities.  

CC Cleaves the junction between single and double-stranded regions o  

CC flap DNA (by similarity).  

CC -COFACTOR: binds 2 magnesium ions per subunit (by similarity).  

CC -SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1  

CC SUBFAMILY.  

CC -----  

CC This SWISS-PROT entry is copyright. It is produced through a collabor  

CC between the Swiss Institute of Bioinformatics and the EMBL outstat  

CC the European Bioinformatics Institute. There are no restrictions on

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch.)

DR EMBL: AJ248285; CAB49654.1; ..
DR PIR: E75117; E75117.
DR HSSP: Q58839; 1A76.
DR HAMAP: MF_00614; .. 1
DR InterPro: IPR000513; EXO_N.1.
DR InterPro: IPR031584; HH2_2.
DR InterPro: IPR006086; XPG_1.
DR InterPro: IPR006085; XPG_N.
DR InterPro: IPR006084; XPG_Rad.
DR Pfam: PF00867; XPG_1; 1.
DR Pfam: PF00752; XPG_N; 1.
DR PRINTS: PR00853; XPGRADSUPER.
DR SMART: SM00279; HH2; 1.
DR SMART: SM00484; XPG1; 1.
DR SMART: SM00485; XPGN; 1.
DR PROSITE: PS00841; XPG_1; FALSE_NEG.
KW Hydrolyase; Nuclease; Endonuclease; Magnesium; Metal-binding;
KW Complete proteome.
FT METAL 154
SQ SEQUENCE 343 AA; 38949 MW; 894A870642B545DD CRC64;

Query Match 32.8%; Score 636; DB 1; Length 343;
Best Local Similarity 41.6%; Pred. No. 2, 9e-35;
Matches 143; Conservative 58; Mismatches 115; Indels 28; Gaps 6;

QY 19 KEKFESEYFGKRIAVDASMSIYFLVVGTEGNETLTNEGEVTSLLQGNFNRITLLEA 78
DB 12 KEELFNLGYCKRIALALCNALYVFLSTIRORDTPLMDSGRITSLSGFLYRTILMLEA 71
QY 79 GIKPVYFEDGKPRDMKKOELAKYRSKRDATKDLTEAVEGDDATLTKSKRIYKTRCH 138
DB 72 GIKPVYFEDGKPRDMKKOELAKYRSKRDATKDLTEAVEGDDATLTKSKRIYKTRCH 131
QY 139 NEDCKRLRLIMGVYVYVAPSEAPAEALCINDKVAVASSEDMSLTFGAPRELRIIMP 198
DB 132 TFOAKKILLOMGPIYCAPSEGEQAQVAVYAGKGDVVASADVDLSLFGTPRZVRRLIT 191
QY 199 SSKKIPVMEVDV-----AVLEELITMQFTDLCGCDY-CISIKIGGCTA 248
DB 192 GKRRKPKGKDIYVYIKPELYLEFVLEKLTPEKLTIELALVCTDYNPGQIKGIPKKA 251
QY 249 KLIRHGSIESLENNKQ---RYQIPEDMPYQEARLKEPMTVDIFELKATPDERG 305
DB 252 EIVK-----YSKDPLAKFQROSDVLYAIKEFFLNPTTDY-SLKKKPHDEE 299
QY 306 LISHLYKNGFNEDRYVYKALEIKTSKNNKSSQGRLESF---KP 346
DB 300 IIRFLCEHDFSEPRVYNGIFELKKAIKAGKQSTLESWFKKKP 343

RESULT 2

FEI_PYROO STANDARD: PPT; 343 AA.
AC OS0123;
DT 26-FEB-2003 (Rel. 41, Created)
DT 26-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Flap structure-specific endonuclease (EC 3.1.1.1).
GN FEN OR PH14.5.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID:53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-OT3;
RX MEDLINE:98344137; PubMed:9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Halkawa Y., Hino Y.,

Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyara A., Nagai Y.,
SA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
SA Furunashi T., Tanaka T., Kudo Y., Yamazaki J., Katsuda N., Gajuchi A.,
SA Aoki K.-I., Yoshizawa Y., Nakamura Y., Robb P.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT *Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998).

CC FUNCTION: Endonuclease that cleave the 5' overhanging flap
CC structure that is generated by displacement synthesis when DNA
CC polymerase encounters the 5' end of a downstream Okazaki fragment.
CC Has 5' endo-/exonuclease and 5' pseudo-Y-endonuclease activities.
CC Cleaves the junction between single and double-stranded regions of
CC flap DNA (By similarity).
CC COFACTOR: Binds 2 magnesium ions per subunit (By similarity).
CC - SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1
CC SUBFAMILY.

EMBL: AP000006; HAA3521.1; ..

DR PIR: A71015; A71015.
DR PDB: 1MC8; 16-OCT-02.
DR HAMAP: MF_00614; .. 1.
DR InterPro: IPR000513; EXO_N.1.
DR InterPro: IPR031584; HH2_2.
DR InterPro: IPR006086; XPG_1.
DR InterPro: IPR006085; XPG_N.
DR InterPro: IPR006084; XPG_Rad.
DR Pfam: PF00867; XPG_1; 1.
DR Pfam: PF00752; XPG_N; 1.
DR PRINTS: PR00853; XPGRADSUPER.
DR SMART: SM00279; HH2; 1.
DR SMART: SM00484; XPG1; 1.
DR SMART: SM00485; XPGN; 1.
DR PROSITE: PS00841; XPG_1; FALSE_NEG.
KW Hydrolyase; Nuclease; Endonuclease; Magnesium; Metal-binding;
KW Complete proteome; 3D-structure.
FT METAL 154
SQ SEQUENCE 343 AA; 38947 MW; 8BFD025F372C1138 CRC64;

Query Match 32.8%; Score 642.5; DB 1; Length 343;
Best Local Similarity 39.2%; Pred. No. 5e-35;
Matches 142; Conservative 66; Mismatches 119; Indels 35; Gaps

QY 1 MCKGLTILADNAKAKKQKFESEYFGKRIAVDASMSIYFLVVGTEGNETLTNEGEVTS 6
DB 1 MGVY ---TGDIVR--KEIDLEN.YCKRIALALCNALYVFLSTIRORDTPLMDSGR 5
QY 63 VTSILQSENNRTILFAGIKPVYFEDGKPRDMKKOELAKYRSKRDATKDLTEAVEGDD 1
DB 54 IISHLSQSEYFTINLHAGIKFAVYFEDGKPRDMKKOELAKYRSKRDATKDLTEAVEGDD 1
QY 121 KDALEKSKPIYKVRQINEDCKRLRLIMGVYVYVAPSEAPAEALCINDKVAVASSED 1
DB 124 LEAKRYQKRAIKVEMILFQAKKLCQIMGPIYCAPSEGEQAQVAVYAGKGDVVASAD 1
QY 191 MCSIIFQAPPEFLIMDSKSKIFVMEVDV-----AVLEELITMQFTDLCGCDY-CIS 2
DB 194 YDSILFQAPPEFLIRNTITGKRRKPKGKDIYVYIKPELYLEFVLEKLTPEKLTIEL 2
QY 232 GTEY-CISIKIGGCTATKIRHGSIESLENNKQ---RYQIPEDMPYQEARLKEPMTVD 2
DB 234 GTDYNPGGVKCIQPKALEIVR-----YSRDPLAKFQROSDVLYAIKEFFLNPT 2
QY 288 NVILDIPELKATPAPEEELISFLVKNQFNEDRYVYKALEIKTSKNNKSSQGRLESF--- 3

DR Pfam: PF01367: 5_3-exonuclease: 1.
 DR Pfam: PF00867: XPG_1: 1.
 DR Pfam: PF00752: XPG_N: 1.
 DR SMART: SM00279: HhH2: 1.
 DR SMART: SM00484: XPG1: 1.
 DR SMART: SM00485: XPGN: 1.
 DR PROSITE: PS00841: XPG_1: 1.
 DR Hydrolyase: Nuclease: Endonuclease: Magnesium: Metal-binding:
 Complete proteome.
 KW METAL.
 FT SEQUENCE 348 AA: 40251 MW: 2058044190087AF9 CRC64.

Query Match 32.1% Score 622: DE 1: Length 348:
 Best Local Similarity 38.9%: Pred. No. 2.5e-34:
 Matches 144: Conservative 73: Mismatches 105: Indels 48: Gaps 12:

QY 1 MGKIGIKTLADNARAKKEKFESEFGRIKIVADSMSTIYQPLVVGTCGKLTINAGE 60
 DB 1 MGKIELRELIFP-----EETDLRALAGRELAIDKFNALYQPLTIKMDG-RPLMDSGR 53
 QY 61 VTSHQGMENRTIRLIFAGIKRPVYVVDCKPPMKQGLAKRYSKDEAVKDLTFAVEYGD 120
 DB 54 ITHLNGLLRYTVNVAEEGIRKPYVFPDEPPLKFTLERRERKEEMKLRRA---KT 110
 QY 121 KDAIEKLSKRYKVRQHNEDCKRLRLMGVPEVAPSEDAEACALCTINDKRVAVSGD 180
 DB 111 KEEREKYARQVARLDESLEVEDAKRLDLMGIPWQAPSEGEACQAYNARCGDVATGSD 170
 QY 181 MDSTFGAPRFLRLHMDPSSKIP-----VMEEDVAKVLEEL-ITDQYIDLC 228
 DB 171 YDLSFGSRILVRNITTYCKRKHPTIGETVYKPRIMLE--DYDQGLSREQLVLA 228
 QY 229 ILGGDY-CDISIKIGOTALKLIRQHS:TESILENLKMDRYQPEMPY-----Q 278
 DB 229 ILGTDYNDGVGPIQPKALCTIRKYSLEDEL-----KD-----IDIMPKIERHLFVPEPE 279
 QY 279 EARLFXEPNVTLDIPELWTPAPDEGLISFLYKONGRNEEDRYTAIKISAKKSSQG 336
 DB 280 KLRFLTEPEVTDYD-ELTMDPEDEEGVETLVEERDSEDRVRAVERLKAEOE:NRKG 338
 QY 339 ---RLSEFP 344
 DB 339 GRCETLDAPF 348

RESULT 13
 FEN_METAC STANDARD: PRT: 338 AA.
 AC 0BT15:
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Flap structure-specific endonuclease (EC 3.1.1.1).
 GN FEN OR MA4004.
 OS Methanosaarcia acetivorans.
 OC Archaea: Euryarchaeota: Euryarchaeia orders incertae sedis:
 CC Methanosaarciales: Methanosaarcinaceae: Methanosaarcina.
 OX NCBI_TaxID=2214:
 RN 11)
 RP SEQUENCE FROM N.A.
 RC STRAIN-GZA / ATCC 35395 / DSM 2634:
 RX MEDLINE=21929760: PubMed=11932218:
 RA Galagan J.E., Nussbaum C., Roy A., Erdtzi M.G., Macdonald P.,
 RA Fitzhugh M., Galvo S., Engels R., Smirnov S., Atwood D., Brown A.,
 RA Allen S., Naylor J., Stange-Thomann N., Deaellano K., Johnson R.,
 RA Linco L., McSwan P., McKernan K., Talamas J., Tjirell A., Ye W.,
 RA Zimmer A., Barber R.D., Cann I., Graham D.E., Graham D.A., Guss A.K.,
 RA Hodderich R., Ingram-Smith C., Kuestner H.C., Krzycki J.A.,
 RA Leish J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
 RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
 RA Perry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
 RA Metcalf W.W., Birren B.:

RT "The genome of Methanosaarcia acetivorans reveals extensive metabolic
 AT and physiological diversity." J.
 RI Genome Res. 12:532-542(2002).
 CC - FUNCTION: Endonuclease that cleave the 5'overhanging flap
 CC structure that is generated by displacement synthesis when DNA
 CC polymerase encounters the 5' end of a downstream Okazaki fragment.
 CC Has 5'endo-/exonuclease and 5'pseudo-Y-endonuclease activities.
 CC Cleaves the junction between single and double-stranded regions of
 CC flap DNA (by similarity).
 CC - COFACTOR: Binds 2 magnesium ions per subunit (by similarity).
 CC - SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1
 CC SUBFAMILY.

CC This SWISS-PRO entry is copyright. It is produced through a collabro
 CC between the Swiss Institute of Bioinformatics and the EMBL outstat
 CC the European Bioinformatics Institute. There are no restrictions on
 CC use by non-profit institutions as long as its context is in no
 CC modified and this statement is not removed. Usage by and for comm
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/seqdb/arcn>
 CC or send an email to license@ebi.ac.uk).

CC EMBL: AEO11112: AAO07354.1: -
 CC HAVAP: MF_00614: -
 CC InterPro: IPR000513: EXO_N.1.
 CC InterPro: IPR001584: HHH_2.
 CC InterPro: IPR006096: XPG_1.
 CC InterPro: IPR006085: XPG_N.
 CC Pfam: PF00867: XPG_1: 1.
 CC Pfam: PF00752: XPG_N: 1.
 CC SMART: SM00279: HhH2: 1.
 CC SMART: SM00484: XPG1: 1.
 CC SMART: SM00485: XPGN: 1.
 CC PROSITE: PS00841: XPG_1: FALSE_NEG.
 CC Hydrolyase: Nuclease: Endonuclease: Magnesium: Metal-binding:
 CC Complete proteome.
 KW METAL.
 FT SEQUENCE 338 AA: 30163 MW: 8523956344F5A67 CRC64.

Query Match 30.2% Score 586.5: DR 1: Length 338:
 Best Local Similarity 38.9%: Pred. No. 5.5e-12:
 Matches 131: Conservative 69: Mismatches 18: Indels 23: Gaps
 QY 18 MKGQKFE SYFGKIVADSMSTIYQPLVVGTCGKLTINAGEVTSHQGMENRTIRL 120
 DB 9 LQKRIKTLSDISNRYAVVADAFNLRLQSLIRQSDSPLVNSKQKVTSHLSGLLYRASL 110
 QY 76 LEAGIKVYVVDCKPPMKQGLAKRYSKDEAVKDLTFAVEYGDQDAIEKLSKRYKVT 180
 DB 69 VVAGIKRVEFPDCKPPLKFTLERRERKEEMKLRRA---KT 110
 QY 121 KQHNEDCKRLRLMGVPEVAPSEDAEACALCTINDKRVAVSGD 180
 DB 111 QVEVEDSKRYLLGIMGIPWQAPSEGEACQAYNARCGDVATGSD 170
 QY 181 MHSKSKIP-----VMEEDVAKVLEELITDQYIDLC 228
 DB 189 AVYCKRKHPTIGETVYKPRIMLE--DYDQGLSREQLVLA 228
 QY 245 QDAIEKLSKRYKVRQHNEDCKRLRLMGVPEVAPSEDAEACALCTINDKRVAVSGD 180
 DB 247 KTAIKIRKHODHIAVR--EKMD---ELEGDQVRLPHEVTDY-FIRWIKPDSE 170
 QY 305 GLISFLYKONGRNEEDRYTAIKISAKKSSQGRLESEF 344
 DB 300 KLIFLLEENLFSIDVYKAKRIRKASNG-ARQKTLDFWF 336

RESULT 11
 FEN_METAC STANDARD: PRT: 351 AA.
 AC 0976H6:
 DT 28-FEB-2003 (Rel. 41, Last sequence update)


```

QY 300 APDEGLISFLKNDGNDRTYKAIEKRS-AKKSSQGLSEFFK 345
DB 280 LPDEKSTIKELVDNFENDFNRVKKHVDKLYLNANKTKOKTIDAMFK 326

RESULT 14
FEN_METMA STANDARD: PRT: 338 AA.
AC 08PRT6:
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Flap structure-specific endonuclease (EC 3.1.1.1).
GN FEN OR MM0906.
OS Methanoscarchina mazel (Methanoscarchina trisla).
OC Archaea; Euryarchaeota; Euryarchaeota orders Incertae sedis;
OC Methanoscarchinales; Methanoscarchinae; Methanoscarchina.
OX NCBI_TaxID=2207;
HN 11
FP SEQUENCE FROM N.A.
RC STRAIN=JG01 / ATCC BAA-159 / DSM 3647 / CCM 88;
RX MEDLINE=22120827; PubMed 12125824;
RA Deppe-Mueller U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA Martinez-Arias R., Henne A., Wietzer A., Baerwald S., Jacobl G.,
RA Bueggemann H., Leonard T., Christmann A., Boemcke M., Stecke R.,
RA Bhattacharya A., Lykidis A., Overbeek R., Klenk H.-F., Gunsalus R.P.,
RA Fitts H.-J., Gottschalk G.;
RT "The genome of Methanoscarchina mazel: evidence for lateral gene
RT transfer between Bacteria and Archaea.";
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
CC 1- FUNCTION: Endonuclease that cleave the 5'overhanging flap
CC structure that is generated by displacement synthesis when DNA
CC polymerase encounters the 5'end of a downstream Okazaki fragment.
CC Has 5'endo-/exonuclease and 5'pseudo-Y-endonuclease activities.
CC Cleaves the junction between single and double-stranded regions of
CC flap DNA (By similarity).
CC 1- COFACTOR: Binds 2 magnesium ions per subunit (By similarity).
CC 1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1
CC SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AE013116; AM30602.1;
DR HAMAP: MF_00614; --1;
DR InterPro: IPR001986; EPSP_synthase.
DR InterPro: IPR000513; EXO_N_H.
DR InterPro: IPR006086; XPG_N.
DR InterPro: IPR006085; XPG_N.
DR InterPro: IPR006084; XPG_N_Rad.
DR Pfam: PF00867; XPG_I; 1.
DR Pfam: PF00752; XPG_N; 1.
DR PRINTS: PR00833; XPGRADSUPER.
DR PROSITE: PS00841; XPG_1; FALSE_NEG.
KW Hydrolyase; Nuclease; Endonuclease; Magnesium; Metal-binding;
KW Complete proteome.
FT METAL 154 154 MAGNESIUM 1 (BY SIMILARITY).
SU SEQUENCE 338 AA; 37711 MW; 59D69E03B0F30E3 CRC64.

Query Match 26.9%; Score 560.5; DB 1; Length 338;
Best Local Similarity 37.3%; Pred. No. 2.9e-30;
Matches 126; Conservative 69; Mismatches 124; Indels 19; Gaps 5;

```

```

QY 16 MKEQKFE--SYFGKIAVDASSTIYQPLIVORTGNETITNEAGVSSHIOGNMRTIRL 75
DB 9 LQPKRIELSDLSNRVAADANTLQFLSTIQRGSPVANSQCKVISHLSGLIYFASL 68
QY 76 LKNGIKPVVFDGKRPDKMKCEILAKRYSKRDADIKDLTAEAVEGDKAIEKLSRYKVY 135

RESULT 15
FEN_METMA STANDARD: PRT: 328 AA.
AC 027670;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Flap structure-specific endonuclease (EC 3.1.1.1).
GN FEN OR MH1633.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales;
OC Methanobacteriales; Methanobacter.
OX NCBI_TaxID=187420;
HN 11
FP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98017514; PubMed 9371463;
RA Smith D.R., Deucelle-Stamm L.A., Belonghery C., Lee H.-M., Dubois J.,
RA Adrege T., Reshizadeh R., Haskaly D., Cook R., Gilbert K.,
RA Hartigan D., Huang L., Kung'u P., Liem M., Rothner F., Gu D.,
RA Shadforth R., Vicario R., Wang Y., Mikolajewski J., Glison R.,
RA Khan N., Gariso A., Bush D., Sater N., Patwell L., Bradshaw S.,
RA McDougall S., Shiner G., Goyal A., Petrovski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling C., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7153(1997).
CC 1- FUNCTION: Endonuclease that cleave the 5'overhanging flap
CC structure that is generated by displacement synthesis when DNA
CC polymerase encounters the 5'end of a downstream Okazaki fragment.
CC Has 5'endo-/exonuclease and 5'pseudo-Y-endonuclease activities.
CC Cleaves the junction between single and double-stranded regions of
CC flap DNA (By similarity).
CC 1- COFACTOR: Binds 2 magnesium ions per subunit (By similarity).
CC 1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1
CC SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collabora
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commerc
CC entities requires a license agreement (See http://www.isb-sib.ch/annou
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AE009222; AAB86106.1;
DR FTR: C69085; C69085.
DR HSSP: O58839; IAT6.
DR HAMAP: MF_00614; --1;
DR InterPro: IPR000513; EXO_N_H.
DR InterPro: IPR000584; XPG_N.
DR InterPro: IPR006086; XPG_I.
DR InterPro: IPR006085; XPG_N.

```

DR InterPro: IP006084; XPGC_Rad.
 DR Pfam: PF00867; XPG_L1;
 DR Pfam: PF00752; XPG_N1;
 DR PRINTS: PR00853; XPGRADSCUPR.
 DR SMART: SM00279; HhH2; 1.
 DR SMART: SM00484; XPG1; 1.
 DR SMART: SM00485; XPGN; 1.
 DR PROSITE: PS00841; XPG_L1; FALSE_NFG.
 DR Hydrolyase: Nuclease; Endonuclease; Magnesium; Metal-binding;
 KM Complete proteome. 154
 FT METAL 154
 SO SEQUENCE 328 AA; 37170 MW; CACIC6046561FDE4 CRC64;

Query Match 28.08; Score 542; DB 1; Length 328;

Best Local Similarity 36.58; Pred. No. 4,7e-29;

Matches 126; Conservative 73; Mismatches 128; Indels 18; Gaps 7;

```

QY 1 MGICGLTKLIAD-NAPKAKKQKFESYFGKKIVADASMSIYQPLIVGRTGMETL:NEAG 59
DB 1 MGVK-----LRDVVSPRR1--RIEDLRGRTVAVDANTLYQFLSS:KQDCTPIKMSRG 52
QY 60 EVTSHLOGMFRRTIRLEAGIKPYVFDGKPPDMKKOEAKRYSKRDDA:KDLTAEVG 119
DB 53 RVTSHLSGILYRTAAVVFRTIRVIVFDGSHHLKGTYSRRADIRKKSLEVFKKALEEG 112
QY 120 OKDAIEKISKRTYKVTROHNEDECKRLRLMGVPVENAPEAEACALCINDVFAVASE 179
DB 113 DIDRAKKYAAVSSHMSSEILESSKRLLELGIPIYQAFCGEAQASYNWKMGDMAVASQ 172
QY 180 DMDSLTFCAPPLHRLMDPSKKIPVMEFDVAVLELFLTMDQFDICIGCCDYCSL 219
DB 173 DYDCLFEGAPRVVKNLTLGKLEDPET-IELESTLRELSTSHQLYDMALLVGTIDFNEGY 211
QY 240 KGIGGOTALKLIROHGSIESILENENKDRYQIPEDWPEQEARLFEKPNVTIDIPETKW 299
DB 232 KGIGARRGLKLIREKGIIFVYIRLENDICGDP-----QYLRIRIFLEPEVSIFY-E:RRR 295
QY 300 APDEEGITFLVKDNGCFNEDRVTKA:EKIKSAKNSQGRLESEF 344
DB 286 KPVEGVYIEFLCTEHEGFSEDRVAALKFPGA--SSYQKSLSDWF 328

```

Search completed: October 22, 2003, 12:40:49
 Job time : 25 secs

DR PROSITE: PS00841: XPG_1: 1
 SQ SEQUENCE 368 AA: 42792 MW: 60146AFA55A73E4 CRC64:
 Query Match 93.2% Score 1807.5: DB 10: Length 360:
 Best Local Similarity 92.1%: Pred. No. 3,2e-14:
 Matches 350: Conservative 16: Mismatches 13: Indels 1: Gaps 1:

QY 1 MSIKGLTKLADNMPKAKKQKESYFGKRIAYVASKSYIOFLIVSGRTGMEILLTNEAGE 60
 |||||||
 DB 1 MGKGLTKLADNMPKAKKQKESYFGKRIAYVASKSYIOFLIVSGRTGMEILLTNEAGE 60
 QY 61 VTSHLQGMENRTIRLLEAGIKPVYVFDGKPPDMKKOELAKRYSKRDATKQCTEAEVGG 120
 |||||||
 DB 61 VTSHLQGMENRTIRLLEAGIKPVYVFDGKPPDMKKOELAKRYSKRDATKQCTEAEVGG 120
 QY 121 KDAIEKSKRYVKTIRQNECCKRLRLMGVVPVADSEAEACALCIDKYVAASD 180
 |||||||
 DB 121 KDAIEKSKRYVKTIRQNECCKRLRLMGVVPVADSEAEACALCIDKYVAASD 180
 QY 181 MDLTFGAPRELRLMDPSSKKIPMEFDVAKVLELELTMDQFIDICILGGDYCDSTK 240
 |||||||
 DB 181 MDLTFGAPRELRLMDPSSKKIPMEFDVAKVLELELTMDQFIDICILGGDYCDSTK 240
 QY 241 GIGGOTALKLRHOSISILENLKORYQIPEDMPYQBARLRFKPNVTLDPPELKTA 300
 |||||||
 DB 241 GIGGOTALKLRHOSISILENLKORYQIPEDMPYQBARLRFKPNVTLDPPELKTA 300
 QY 301 PDEGLISFLYKDNQFNEDRYTKAIEKISAKNKSOGRCJFSEFKPATISAPLKPKETS 360
 |||||||
 DB 301 PDEGLISFLYKDNQFNEDRYTKAIEKISAKNKSOGRCJFSEFKPATISAPLKPKETS 360
 QY 361 DKTSTAAANKKTK-AGGKK 379
 |||||||
 DB 361 FKPTKAVANKTKTGAGKK 380

RESULT 2
 ID 065251 PRELIMINARY: PRT: 362 AA.
 AC 065251;
 DT 01-AUG-1998 (TREMBLrel: 07: Created)
 DT 01-AUG-1998 (TREMBLrel: 07: Last sequence update)
 DT 01-OCT-2002 (TREMBLrel: 22: Last annotation update)
 DE F2IE-0.3 protein.
 GN F2IE0.3
 OS Arabidopsis thaliana (Mouse-ear cress);
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots; Rosidae;
 OC Eurosidia II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Davidson S., Kolling T., David M., O'Brien D.;
 RT The sequence of A. thaliana F2IE0.3;
 RL Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.
 RN 12
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.
 RN 13
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.
 RN 14
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF058914: AAC13596.1: -

DR HSRP: Q58839: 1A76.
 DR InterPro: IPR002421: 5_3_exonuclease.
 DR InterPro: IPR000513: Exo_N_1.
 DR InterPro: IPR003584: HH2.
 DR InterPro: IPR006984: XPGC_Pad.
 DR InterPro: IPR006086: XPG_I.
 DR InterPro: IPR006095: XPG_N.
 DR Pfam: PF01957: 5_3_exonuclease_1.
 DR Pfam: PF00567: XPG_1.
 DR Pfam: PF00752: XPG_N_1.
 DR PRINTS: PR00853: XPGRADSUPER.
 DR SMART: SM00279: H332: 1.
 DR SMART: SM00484: XPGC: 1.
 DR SMART: SM00485: XPGN: 1.
 DR PROSITE: PS00841: XPG_1: 1.
 SQ SEQUENCE 362 AA: 41205 MW: A4DD170C3AD25D CRC64:
 Query Match 67.7% Score 1313.5: DB 10: Length 362:
 Best Local Similarity 72.1%: Pred. No. 4,2e-88:
 Matches 259: Conservative 30: Mismatches 33: Indels 37: Gaps

QY 1 MGKGLTKLADNMPKAKKQKESYFGKRIAYVASKSYIOFLIVSGRTGMEILLTNEAGE 60
 |||||||
 DB 1 MGKGLTKLADNMPKAKKQKESYFGKRIAYVASKSYIOFLIVSGRTGMEILLTNEAGE 60
 QY 61 VTSHLQGMENRTIRLLEAGIKPVYVFDGKPPDMKKOELAKRYSKRDATKQCTEAEVGG 120
 |||||||
 DB 61 VTSHLQGMENRTIRLLEAGIKPVYVFDGKPPDMKKOELAKRYSKRDATKQCTEAEVGG 120
 QY 121 KDAIEKSKRYVKTIRQNECCKRLRLMGVVPVADSEAEACALCIDKYVAASD 180
 |||||||
 DB 121 KDAIEKSKRYVKTIRQNECCKRLRLMGVVPVADSEAEACALCIDKYVAASD 180
 QY 181 MDLTFGAPRELRLMDPSSKKIPMEFDVAKVLELELTMDQFIDICILGGDYCDSTK 240
 |||||||
 DB 181 MDLTFGAPRELRLMDPSSKKIPMEFDVAKVLELELTMDQFIDICILGGDYCDSTK 240
 QY 241 GIGGOTALKLRHOSISILENLKORYQIPEDMPYQBARLRFKPNVTLDPPELKTA 300
 |||||||
 DB 241 GIGGOTALKLRHOSISILENLKORYQIPEDMPYQBARLRFKPNVTLDPPELKTA 300
 QY 301 PDEGLISFLYKDNQFNEDRYTKAIEKISAKNKSOGRCJFSEFKPATISAPLKPKETS 360
 |||||||
 DB 301 PDEGLISFLYKDNQFNEDRYTKAIEKISAKNKSOGRCJFSEFKPATISAPLKPKETS 360
 QY 361 DKTSTAAANKKTK-AGGKK 379
 |||||||
 DB 361 FKPTKAVANKTKTGAGKK 380

RESULT 3
 ID 070043 PRELIMINARY: PRT: 382 AA.
 AC 070043;
 DT 01-FEB-1997 (TREMBLrel: 02: Created)
 DT 01-FEB-1997 (TREMBLrel: 02: Last sequence update)
 DT 01-OCT-2002 (TREMBLrel: 22: Last annotation update)
 DE 5' nuclease XEN-A.
 GN XEN-A.
 OS Xenopus laevis (African clawed frog);
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
 OC Xenopodidae; Xenopus.
 OX NCBI_TaxID=8355;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Cox L.S.;
 RA Bibikova M., Chi E., Wu B., Kim K., Carroll J.;
 RT Cloning and investigation of Xenopus Fenu: developmental expression
 and function in DNA replication;
 RL Submitted (May-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U64563: AA006176.1: -
 DR EMBL: AF055397: AA002814.1: -
 DR HSRP: Q58839: 1A76.

DR InterPro: IPR002421; 5_3_exonuclease.
 DR InterPro: IPR000513; Exo_N_1.
 DR InterPro: IPR003584; HNH_2.
 DR InterPro: IPR006084; XPG_Rad.
 DR InterPro: IPR006086; XPG_I.
 DR InterPro: IPR006085; XPG_N.
 DR Pfam: PF01367; 5_3_exonuclease; 1.
 DR Pfam: PF00867; XPG_1; 1.
 DR Pfam: PF00752; XPG_N; 1.
 DR PRINTS: PR00853; XPGRADSEPER.
 DR SMART: SMO0279; HNH2; 1.
 DR SMART: SMO0484; XPGI; 1.
 DR SMART: SMO0485; XPGN; 1.
 DR Endonuclease.
 DR SEQUENCE: 382 AA; 42668 MW; 981DB0EDAD158D57 CRC64;

Query Match 55.1%; Score 1069; DR 13; Length 382;
 Best Local Similarity 55.7%; Pred. No. 3,6e-70;
 Matches 210; Conservative 63; Mismatches 98; Indels 6; Gaps 4;

QY 1 MGIGKGLKLLADNAPKAKKEQKFESEYFGKILAVDASMSIYQFLIVSGTGMTLTNAGE 63
 D 1 MGIGGLAKLLADNAPAKIKENDIKSYFGKAVVADSMCTIYQFLIARVDG-NITQNEGE 59
 QY 61 VTSILQGMENRTIRLEAGIKPVYVFDCKPPMKKQELAKRYSKDCTKDLTEAVEGQ 120
 D 60 TSHLMGMFYRTIRIEMLEHGIKPVYVFDCKPPMKSGELAKSRRAEAKILLEAEAGE 119
 QY 121 KDAIEKLSKRYKVTQHNEDCKRLRLMGVAVVAPSEAECAALCINDKRVAVASED 180
 D 120 VENIEKTRKLVKVKYKQHNEDCKRLTLTNGIYVADPEAEKTCALVAKKAVVAAVED 179
 QY 181 MSITFGARFLRLHLDPSKKIPWNEFDVAKVLELEITMDQFLDCLTCCDVCSTK 240
 D 180 MDATFGTIVLLRLHLTASAKKLIPIQFHLNIVDIDIGITIHQFVLDLILSSDYCEIR 239
 QY 241 GIGGQIALKLIROHGSISLENLKDRYQIPEDMPYQARLRLEPNVT-LDIPREKWT 299
 D 240 GIGPKAIDLIRQHKIDIELINIDIKKYPENMLHKEAKLLEPEVVTDLITLKW 298
 QY 300 APDEEGISFLVQDNFNEEDRYTKATEIKSAKNSQOCLESFEPKPTATSAFIRKET 359
 D 300 EPDEEGIVAFMCGEKEKFSEDRIKNGAKKAKKNGOSTQGRDIDPFKVGISVS-TKRKA 358
 QY 360 SEKTSKAANKKIRKAG 376
 D 359 E---SKGSAKKRAKTGG 372

RESULT 4

ID P70054 PRELIMINARY; PRI: 382 AA.
 AC P70054;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE XENB.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;
 OC Xenopodinae; Xenopus.
 CC NCBI_TaxID=8355;
 XX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Head;
 RA MEDLIN:99059415; PubMed=9852084;
 RA Bidikova M., Wu B., Chl E., Kim K.H., Trautman J.K., Carroll D.;
 RT "Characterization of FEN-1 from Xenopus laevis. cDNA cloning and role
 in DNA metabolism."
 RT J. Biol. Chem. 273:34222-34229(1998).
 RI EMBL: U68141; AAB08478.1;
 DR HSP: Q58833; IAT6;
 DR InterPro: IPR002421; 5_3_exonuclease.
 DR InterPro: IPR002421; 5_3_exonuclease.

DR InterPro: IPR000513; Exo_N_1.
 DR InterPro: IPR003584; HNH_2.
 DR InterPro: IPR006084; XPG_Rad.
 DR InterPro: IPR006086; XPG_I.
 DR InterPro: IPR006085; XPG_N.
 DR Pfam: PF01367; 5_3_exonuclease; 1.
 DR Pfam: PF00867; XPG_1; 1.
 DR Pfam: PF00752; XPG_N; 1.
 DR PRINTS: PR00853; XPGRADSEPER.
 DR SMART: SMO0279; HNH2; 1.
 DR SMART: SMO0484; XPGI; 1.
 DR SMART: SMO0485; XPGN; 1.
 DR SEQUENCE: 382 AA; 42865 MW; 10C48916A232C46C CRC64;

Query Match 55.1%; Score 1068; DR 13; Length 382;
 Best Local Similarity 56.2%; Pred. No. 4.2e-70;
 Matches 222; Conservative 60; Mismatches 99; Indels 6; Gaps 4;

QY 1 MGIGKGLKLLADNAPKAKKEQKFESEYFGKILAVDASMSIYQFLIVSGTGMTLTNAGE 63
 D 1 MGIGGLAKLLADNAPAKIKENDIKSYFGKAVVADSMCTIYQFLIARVDG-NITQNEGE 59
 QY 61 VTSILQGMENRTIRLEAGIKPVYVFDCKPPMKKQELAKRYSKDCTKDLTEAVEGQ 120
 D 60 TSHLMGMFYRTIRIEMLEHGIKPVYVFDCKPPMKSGELAKSRRAEAKILLEAEAGE 119
 QY 121 KDAIEKLSKRYKVTQHNEDCKRLRLMGVAVVAPSEAECAALCINDKRVAVASED 180
 D 120 VENIEKTRKLVKVKYKQHNEDCKRLTLTNGIYVADPEAEKTCALVAKKAVVAAVED 179
 QY 181 MSITFGARFLRLHLDPSKKIPWNEFDVAKVLELEITMDQFLDCLTCCDVCSTK 240
 D 180 MDATFGTIVLLRLHLTASAKKLIPIQFHLNIVDIDIGITIHQFVLDLILSSDYCEIR 239
 QY 241 GIGGQIALKLIROHGSISLENLKDRYQIPEDMPYQARLRLEPNVT-LDIPREKWT 299
 D 240 GIGPKAIDLIRQHKIDIELINIDIKKYPENMLHKEAKLLEPEVVTDLITLKW 298
 QY 300 APDEEGISFLVQDNFNEEDRYTKATEIKSAKNSQOCLESFEPKPTATSAFIRKET 359
 D 300 EPDEEGIVAFMCGEKEKFSIRIRINAKKLAARQGSTQGRDIDPFKVGISVS-TKRKA 358
 QY 360 SEKTSKAANKKIRKAG 376
 D 359 E---SKSTKRSKTSKG 372

RESULT 5

ID O57351 PRELIMINARY; PRI: 382 AA.
 AC O57351;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Flap endonuclease 1.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;
 OC Xenopodinae; Xenopus.
 CC NCBI_TaxID=8355;
 XX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE:98204872; PubMed=9855864;
 RA Kim K., Blade S., Matsumoto Y.;
 RT "Involvement of Flap endonuclease 1 in base excision DNA repair."
 RT J. Biol. Chem. 273:8542-8546(1998).
 DR EMBL: AF036327; AAB84707.1;
 DR HSP: Q58833; IAT6;
 DR InterPro: IPR002421; 5_3_exonuclease.
 DR InterPro: IPR000513; Exo_N_1.
 DR InterPro: IPR003584; HNH_2.
 DR InterPro: IPR006084; XPG_Rad.

DR EMBL: AK078778: BAC37390.1: -
 DR MGD: MGI:102779: Pen1
 DR InterPro: IPR002421: 5_3-exonuclease.
 DR InterPro: IPR000513: Exo_N1.
 DR InterPro: IPR003584: HHH_2.
 DR InterPro: IPR006084: XPGC_Rad.
 DR InterPro: IPR006086: XPG_I.
 DR InterPro: IPR006085: XPG_N.
 DR Pfam: PF01367: 5_3-exonuclease: 1.
 DR Pfam: PF00867: XPG_I: 1.
 DR Pfam: PF00752: XPG_N: 1.
 DR PRINTS: PR00853: XPGRADSUPER.
 DR SMART: SM00279: HNH2: 1.
 DR SMART: SM00484: XPG1: 1.
 DR SMART: SM00485: XPGN: 1.
 DR PROSITE: PS00841: XPG_1: 1.
 DR PROSITE: PS00842: XPG_2: 1.
 DR Endonuclease.
 KW SEQUENCE 380 AA: 42623 MW: 185903288046520D CRC64:

Query Match 53.8% Score 1043: DB 11: Length 380:
 Best Local Similarity 53.8% Pred. No. 2.8e-68:
 Matches 205: Conservative 69: Mismatches 99: Indels 8: Gaps 5:

QY 1 MGIGLTKLLADNAPKAMKQKFESEYFGKRIADVADSMSTYQFLIVGRTGMEITLNEAGE 60
 DB 1 MGIGLTKLLADNAPKAMKQKFESEYFGKRIADVADSMSTYQFLIVGRTGMEITLNEAGE 60
 DB 2 MGIGLTKLLADNAPKAMKQKFESEYFGKRIADVADSMSTYQFLIVGRTGMEITLNEAGE 59
 QY 61 VTSILGMEFNTIRLEAGIKPVYVFDGPPPMKKOELAKRSKDDATKDLTEAVFVG 120
 DB 60 TTSILGMEFNTIRLEAGIKPVYVFDGPPPMKKOELAKRSKDDATKDLTEAVFVG 119
 DB 60 TTSILGMEFNTIRLEAGIKPVYVFDGPPPMKKOELAKRSKDDATKDLTEAVFVG 119
 QY 121 KDAIEKLSRTYKVTIQHNEDECKLRRLMGVPEVAPSEAEAFCAALCINKEVAVASD 160
 DB 120 EEEVEKFKRLKVKVTKQHNDCKHLISLNGIPYLAPSEAEAFCAALCINKEVAVASD 159
 DB 120 EEEVEKFKRLKVKVTKQHNDCKHLISLNGIPYLAPSEAEAFCAALCINKEVAVASD 159
 QY 181 MSLIFGAPRFLHLMDSKKIPVMEPVAVKVEELHLMQFIDLCILGCGDVCISIK 240
 DB 180 MSLIFGAPRFLHLMDSKKIPVMEPVAVKVEELHLMQFIDLCILGCGDVCISIK 239
 DB 180 MSLIFGAPRFLHLMDSKKIPVMEPVAVKVEELHLMQFIDLCILGCGDVCISIK 239
 QY 241 GIGGCTALKLIRHQSIESILENLNDRYQIPEDPYQAFRLFKPNTLD--IPKWK 298
 DB 240 GIGGCTALKLIRHQSIESILENLNDRYQIPEDPYQAFRLFKPNTLD--IPKWK 298
 DB 240 GIGGCTALKLIRHQSIESILENLNDRYQIPEDPYQAFRLFKPNTLD--IPKWK 298
 QY 249 TAPDEGLISFLVKNDFNEEDVTKAIEKISAKNKSOGRISEFFKPTATTSAPLAKKE 358
 DB 249 TAPDEGLISFLVKNDFNEEDVTKAIEKISAKNKSOGRISEFFKPTATTSAPLAKKE 357
 DB 249 TAPDEGLISFLVKNDFNEEDVTKAIEKISAKNKSOGRISEFFKPTATTSAPLAKKE 357
 QY 359 TSDKTSKAANKTKAGKKK 379
 DB 358 TSDKTSKAANKTKAGKKK 375

RESULT 8

Q8C5X6

ID Q8C5X6

AC Q8C5X6

PRELIMINARY: PRT: 411 AA.

DT 01-MAR-2003 (ITEMBLRE). 23, Created)

DT 01-MAR-2003 (ITEMBLRE). 23, Last sequence update)

DT 01-MAR-2003 (ITEMBLRE). 23, Last annotation update)

DF 54ap structure specific endonuclease 1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI_TaxID=10090;

KN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=2234683; PubMed=1246851;
 RA The FANIM Consortium.
 RA The FANIM Genome Exploration Research Group Phase I & II Team;
 RA "Analysis of the mouse transcriptome based on functional annotation of
 65,770 full-length cDNAs."

Hi: Nature 420:563-573(2002).
 DR EMBL: AK078777: BAC36544.1: -
 DR SEQUENCE 411 AA: 45819 MW: 00FB31520454032D CRC64:

Query Match 53.8% Score 1043: DB 11: Length 411:
 Best Local Similarity 53.8% Pred. No. 3.2e-68:
 Matches 205: Conservative 69: Mismatches 99: Indels 8: Gaps 5:

QY 1 MGIGLTKLLADNAPKAMKQKFESEYFGKRIADVADSMSTYQFLIVGRTGMEITLNEAGE 60
 DB 1 MGIGLTKLLADNAPKAMKQKFESEYFGKRIADVADSMSTYQFLIVGRTGMEITLNEAGE 60
 DB 1 MGIGLTKLLADNAPKAMKQKFESEYFGKRIADVADSMSTYQFLIVGRTGMEITLNEAGE 59
 QY 61 VTSILGMEFNTIRLEAGIKPVYVFDGPPPMKKOELAKRSKDDATKDLTEAVFVG 120
 DB 60 TTSILGMEFNTIRLEAGIKPVYVFDGPPPMKKOELAKRSKDDATKDLTEAVFVG 119
 DB 60 TTSILGMEFNTIRLEAGIKPVYVFDGPPPMKKOELAKRSKDDATKDLTEAVFVG 119
 QY 121 KDAIEKLSRTYKVTIQHNEDECKLRRLMGVPEVAPSEAEAFCAALCINKEVAVASD 160
 DB 120 EEEVEKFKRLKVKVTKQHNDCKHLISLNGIPYLAPSEAEAFCAALCINKEVAVASD 159
 DB 120 EEEVEKFKRLKVKVTKQHNDCKHLISLNGIPYLAPSEAEAFCAALCINKEVAVASD 159
 QY 181 MSLIFGAPRFLHLMDSKKIPVMEPVAVKVEELHLMQFIDLCILGCGDVCISIK 240
 DB 180 MSLIFGAPRFLHLMDSKKIPVMEPVAVKVEELHLMQFIDLCILGCGDVCISIK 239
 DB 180 MSLIFGAPRFLHLMDSKKIPVMEPVAVKVEELHLMQFIDLCILGCGDVCISIK 239
 QY 241 GIGGCTALKLIRHQSIESILENLNDRYQIPEDPYQAFRLFKPNTLD--IPKWK 298
 DB 240 GIGGCTALKLIRHQSIESILENLNDRYQIPEDPYQAFRLFKPNTLD--IPKWK 298
 DB 240 GIGGCTALKLIRHQSIESILENLNDRYQIPEDPYQAFRLFKPNTLD--IPKWK 298
 QY 249 TAPDEGLISFLVKNDFNEEDVTKAIEKISAKNKSOGRISEFFKPTATTSAPLAKKE 358
 DB 249 TAPDEGLISFLVKNDFNEEDVTKAIEKISAKNKSOGRISEFFKPTATTSAPLAKKE 357
 DB 249 TAPDEGLISFLVKNDFNEEDVTKAIEKISAKNKSOGRISEFFKPTATTSAPLAKKE 357
 QY 359 TSDKTSKAANKTKAGKKK 379
 DB 358 TSDKTSKAANKTKAGKKK 375

RESULT 9

Q8R069

ID Q8R069

AC Q8R069

PRELIMINARY: PRT: 380 AA.

DT 01-JUN-2002 (ITEMBLRE). 21, Created)

DT 01-JUN-2002 (ITEMBLRE). 21, Last sequence update)

DT 01-MAR-2003 (ITEMBLRE). 23, Last annotation update)

DE Hypothetical 42.6 kDa protein.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

KN 11
 RP SEQUENCE FROM N.A.
 RC Strausberg R.
 RC Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC027295: AAH27295.1: -
 DR InterPro: IPR000513: Exo_N1.
 DR InterPro: IPR003584: HHH_2.
 DR InterPro: IPR006086: XPG_I.
 DR InterPro: IPR006085: XPG_N.
 DR Pfam: PF00867: XPG_I: 1.
 DR Pfam: PF00752: XPG_N: 1.
 DR SMART: SM00279: HNH2: 1.
 DR SMART: SM00484: XPG1: 1.
 DR SMART: SM00485: XPGN: 1.
 DR PROSITE: PS00841: XPG_1: 1.
 DR PROSITE: PS00842: XPG_2: 1.
 DR Hypothetical protein.
 KW SEQUENCE 380 AA: 42609 MW: CBB38745358F55AF CRC64:

Query Match 53.7% Score 1042: DB 11: Length 380:
 Best Local Similarity 53.8% Pred. No. 3.4e-68:
 Matches 204: Conservative 70: Mismatches 99: Indels 8: Gaps 5:

QY 1 MGIGLTKLLADNAPKAMKQKFESEYFGKRIADVADSMSTYQFLIVGRTGMEITLNEAGE 60

```

DQ 1 MGTHGLAKLLADVAPSAIRENDKSYFGKRVKALDASMSYQFLIAT-KGGGVYQNEBGF 59
QY 61 VTSHLQGMFNRIITLLEAGIKPVYVFDGKPPROKKQOELAKRSYKDDACTKDLTEAVEVGD 120
DB 60 TTSHLMGMFYRTIRKMGNGVYKPVYVFDGKPPOLKSGELAKRSERBRAEAKKQIQOAOEAG 119
QY 121 KDAIEKLSKRTVAVTQHNEDCKRLRLKMGVYVFAEFAEACALCINDVAFVAVASD 180
DB 120 EEVEKFTKRLVAVTQHNEDCKRLRLKMGVYVFAEFAEACALCINDVAFVAVASD 179
QY 181 MDLTFGAPRFLRLHMDPSSSKIPVMEFDVAKVLELELMDQFIDICTCCGDCYCSIK 240
DB 180 MDCLTFGSPVLMRHLTASFAKKLPIQEFHLSRVLOELGMOEQFVDCILIGSDYCESIR 239
QY 241 GIGGOTALKLIRQGSSTLELENMKDRIOIPEDMPYQEARLRFKEPNVTL--IPELKW 299
DB 240 GIGKRAVDLTOKHKSIEETVRRLDPSKYPVPEMMLHKEAQQFLPEVY-LDPESVELKW 299
QY 299 TAPDEGLISFLVNDGNEFDVATKAIKIKSAKNSOGRLSEFKEPTATSAFLKKE 358
DB 299 SEPDEELVKEFGCKHGFSEERIRSGVKRLSKSGOSTGRLDDPFTKVTGSLSS-AKKE 357
QY 359 TSDKTSKAANKKTKAGGKK 379
DB 358 PE--PKGPAKKKAKTGAGCK 375

```

RESULT 10

```

QY 09JHM7 PRELIMINARY: PRI: 380 AA.
AC 09JHM7
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Flag structure-specific endonuclease.
CN FEN-1
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen D., Cao G., Yang S., Li M., Chen J.;
RT "Cloning and characterization of a rat DNA structure-specific
RT endonuclease (Fen-1).";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF281018; AAF81265.1;
DR HSSP: Q58839; IAT6.
DR InterPro: IPR002423; 5.3_exonuclease.
DR InterPro: IPR000513; Fxo_N_Lf.
DR InterPro: IPR003584; HNH_2.
DR InterPro: IPR006084; XPG_Rad.
DR InterPro: IPR006085; XPG_N.
DR Pfam: PF01367; 5_3_exonuclease;
DR Pfam: PF00867; XPG_I;
DR Pfam: PF00752; XPG_N;
DR PRINTS: PR00853; XTRADSUPER.
DR SMART: SM00279; HNH2;
DR SMART: SK00484; XPG1;
DR SMART: SK00485; XPGN;
DR PROSITE: PS00842; XPG_2;
KM Endonuclease.
SQ
SEQUENCE 380 AA: 42622 MW: 61698CET16F182136 CRC64;

```

Query Match

Best Local Similarity 53.6%; Score 1040; DB 11; Length 380;
Matches 200; Conservative 73; Mismatches 101; Indels 6; Gaps 4;

```

QY 1 MGIGLTKLLADNPKAMKEQKESYFGKRIADVASTIYQFLIVGRIGTCEITLNEAG 60
DB 1 MGTHGLAKLLADVAPSAIRENDKSYFGKRVKALDASMSYQFLIAT-KGGGVYQNEBGF 59

```

```

QY 61 VTSHLQGMFNRIITLLEAGIKPVYVFDGKPPROKKQOELAKRSYKDDACTKDLTEAVEVGD 120
DB 60 TTSHLMGMFYRTIRKMGNGVYKPVYVFDGKPPOLKSGELAKRSERBRAEAKKQIQOAOEAG 119
QY 121 KDAIEKLSKRTVAVTQHNEDCKRLRLKMGVYVFAEFAEACALCINDVAFVAVASD 180
DB 120 EEVEKFTKRLVAVTQHNEDCKRLRLKMGVYVFAEFAEACALCINDVAFVAVASD 179
QY 181 MDLTFGAPRFLRLHMDPSSSKIPVMEFDVAKVLELELMDQFIDICTCCGDCYCSIK 240
DB 180 MDCLTFGSPVLMRHLTASFAKKLPIQEFHLSRVLOELGMOEQFVDCILIGSDYCESIR 239
QY 241 GIGGOTALKLIRQGSSTLELENMKDRIOIPEDMPYQEARLRFKEPNVTL--IPELKW 299
DB 240 GIGKRAVDLTOKHKSIEETVRRLDPSKYPVPEMMLHKEAQQFLPEVY-LDPESVELKW 299
QY 299 TAPDEGLISFLVNDGNEFDVATKAIKIKSAKNSOGRLSEFKEPTATSAFLKKE 358
DB 299 SEPDEELVKEFGCKHGFSEERIRSGVKRLSKSGOSTGRLDDPFTKVTGSLSS-AKKE 357
QY 359 TSDKTSKAANKKTKAGGKK 379
DB 358 PE--PKGPAKKKAKTGAGCK 375

```

RESULT 11

```

QY 08C952 PRELIMINARY: PRI: 380 AA.
AC 08C952
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Flag structure specific endonuclease 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-C57BL/6J; Tissue=Cerebellum;
RC MEDLINE-22354683; PubMed-12465851;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:565-573(2002);
RL EMBL: AK042942; BAC031452.1;
SQ
SEQUENCE 380 AA: 42589 MW: 63488B271859686E CRC64;

```

Query Match

Best Local Similarity 53.5%; Score 1037; DB 11; Length 380;
Matches 204; Conservative 69; Mismatches 100; Indels 8; Gaps

```

QY 1 MGIGLTKLLADNPKAMKEQKESYFGKRIADVASTIYQFLIVGRIGTCEITLNEAG 60
DB 1 MGTHGLAKLLADVAPSAIRENDKSYFGKRVKALDASMSYQFLIAT-KGGGVYQNEBGF 59
QY 61 VTSHLQGMFNRIITLLEAGIKPVYVFDGKPPROKKQOELAKRSYKDDACTKDLTEAVEVGD 120
DB 60 TTSHLMGMFYRTIRKMGNGVYKPVYVFDGKPPOLKSGELAKRSERBRAEAKKQIQOAOEAG 119
QY 121 KDAIEKLSKRTVAVTQHNEDCKRLRLKMGVYVFAEFAEACALCINDVAFVAVASD 180
DB 120 EEVEKFTKRLVAVTQHNEDCKRLRLKMGVYVFAEFAEACALCINDVAFVAVASD 179
QY 181 MDLTFGAPRFLRLHMDPSSSKIPVMEFDVAKVLELELMDQFIDICTCCGDCYCSIK 240
DB 180 MDCLTFGSPVLMRHLTASFAKKLPIQEFHLSRVLOELGMOEQFVDCILIGSDYCESIR 239
QY 241 GIGGOTALKLIRQGSSTLELENMKDRIOIPEDMPYQEARLRFKEPNVTL--IPELKW 299
DB 240 GIGKRAVDLTOKHKSIEETVRRLDPSKYPVPEMMLHKEAQQFLPEVY-LDPESVELKW 299

```

QY 259 TAPDEGLISFLVKNDFNEDRVTKATEIKSAKNSGSGHLESEFKPTATISAPLKKRK 356
 DB 299 SEPNEELFKVMCGKQFSEERIRSGVKRLSKSGSTGRLDQFVKVTSLSL-ARRKE 357
 QY 359 TSDKTSKAANKKTKAGSKK 379
 DB 358 PE---PKCPAKKAKTGAGK 375

RESULT 12

ID Q9N3T2 PRELIMINARY: PRT: 382 AA.
 AC Q9N3T2;
 DT 01-OCT-2000 (TrEMBL: 15, Created)
 DT 01-OCT-2003 (TrEMBL: 15, Last sequence update)
 DT 01-MAR-2003 (TrEMBL: 23, Last annotation update)
 DE Hypothetical 42.5 kDa protein.
 GN Y4766A.8.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematozoa; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodolerinae; Caenorhabditis.
 CX NCBI_TaxID:6239;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=93069613; PubMed=9851916;
 RA Note:
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium."
 RT Science 282:2012-2018 (1998).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Minx P., Graves T., Hawrysko C.;
 RT "The sequence of C. elegans cosmid Y4766A."
 RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RT "Direct Submission."
 RT Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC024751; AAF60553.1;
 DR HSSP: Q58839; 1A76.
 DR Wormpep: Y4766A.8; CD22109.
 DR InterPro: IPR002421; 5_3_exonuclease.
 DR InterPro: IPR000513; Exo_N_1.
 DR InterPro: IPR003584; HHH_2.
 DR InterPro: IPR006084; XPG_Rad.
 DR InterPro: IPR006086; XPG_I.
 DR InterPro: IPR006085; XPG_N.
 DR Pfam: PF01367; 5_3_exonuclease; 1.
 DR Pfam: PF00867; XPG_1; 1.
 DR Pfam: PF00752; XPG_N_1; 1.
 DR PRINTS: PR00853; XPGRADSUPR.
 DR SMART: SM00279; HHH2; 1.
 DR SMART: SM00484; XPGI; 1.
 DR SMART: SM00485; XPGI; 1.
 DR SMART: SM00485; XPGN; 1.
 DR SMART: PS00842; XPG_2; 1.
 DR Hypothetical protein.
 KW SPONDICP 382 AA; MW: 42549 MW; FDI:0BA0707321EB CRC64;

Query Match: 53.5%; Score 1036.5; DB 5; Length 382;
 Best Local Similarity: 52.3%; Pred. No. 8,66-68;
 Matches 201; Conservative 73; Mismatches 99; Indels 11; Gaps 4;

QY 1 MGKGLTKLIADNAKAKKQKFESEYFGRKLAVDASSTYFLLVCGTCKEHTITNAGE 60
 DB 1 MGKGLTSOVIADNAKSAIRKVENKAFGRVIAVDSKCTYOFLLAVRQDSQ-LGSEEGE 59
 QY 51 VTSHQGPNKRTIRLEAGIKPVYVFGKPPDMKROGCAKRYSKRDYATKLLIYAVVGD 120
 DB 60 TISILMKIKRTYRKFENGVKVFYVDFUKYDPMKGGLERKSERRAEKALIFAKKRD 119

QY 121 KDAIFKLSKRPVYTRQENEDKFLLR:MGVYVAFAPSAFAPCAACINDKVFVAVASID
 DB 120 VKAEKFERRLKVKTKQONIPAKKRLGIMGIVVEAPEVAPCAQHVKAKKVFQVTEID
 QY 181 MSLTFCGAPRPIRLMDPSKRIIVMEPDVAKVLEETFLMDQFIDICIGQDYDSIK
 DB 180 MDALFTSTVLCRHLFAPVAKKIKTKERNLSIALBKKSVEEELICILCGQYCTIR
 QY 241 GIGGOTAKLIRQSGSISLESLANKDPIYCIPEPWPTQEARLKEFNVL-DIPELKWIT
 DB 240 GVGPKKAVELIKQHNIEITLLENLQNKVPPEDMPYKRAHLEFLNEPVTKPEVEELTWK
 QY 300 APDEGLISFLVKNDFNEDRVTKATEIKSAKNSGSGHLESEFKPTATISAPLKKRK 356
 DB 300 VADVKGAVLOFLGQENFNEERKMAIAKIKTSKSGSTGRLDQFVKVTSLSL-ARRKE 357
 QY 359 TSDKTSKAANKKTKAGSKK 379
 DB 360 KAEKAKKAKKA---KKQGRK 378

RESULT 13

ID Q9N3T2 PRELIMINARY: PRT: 378 AA.
 AC Q9N3T2;
 DT 01-JUN-2001 (TrEMBL: 17, Created)
 DT 01-JUN-2001 (TrEMBL: 17, Last sequence update)
 DT 01-OCT-2002 (TrEMBL: 22, Last annotation update)
 DE Hyp endonuclease-1.
 GN FEN1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxID:10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=129SvEv;
 RA KaranJawala N.E., Shi X., Hsieh C.-T., Fierke M.R.;
 RT "The Mammalian Fen1 Locus: Structure and Conserved Sequence
 RT Features."
 RT Microb. Comp. Genomics 0:0-0 (2001).
 DR EMBL: AY014962; AAK01853.1;
 DR HSSP: Q58839; 1A76.
 DR InterPro: IPR002421; 5_3_exonuclease.
 DR InterPro: IPR000513; Exo_N_1.
 DR InterPro: IPR003584; HHH_2.
 DR InterPro: IPR006084; XPG_Rad.
 DR InterPro: IPR006086; XPG_I.
 DR InterPro: IPR006085; XPG_N.
 DR Pfam: PF01367; 5_3_exonuclease; 1.
 DR Pfam: PF00867; XPG_1; 1.
 DR Pfam: PF00752; XPG_N_1; 1.
 DR PRINTS: PR00853; XPGRADSUPR.
 DR SMART: SM00279; HHH2; 1.
 DR SMART: SM00464; XPGI; 1.
 DR SMART: SM00485; XPGN; 1.
 DR PROSITE: PS00841; XPG_1; 1.
 DR PROSITE: PS00842; XPG_2; 1.
 DR Endonuclease.
 KW SPONDICP 378 AA; MW: 827940ABACAC9P39 CRC64;

Query Match: 52.0%; Score 1009; DB 11; Length 378;
 Best Local Similarity: 52.9%; Pred. No. 8,70-66;
 Matches 221; Conservative 69; Mismatches 102; Indels 8; Gaps

QY 1 MGKGLTKLIADNAKAKKQKFESEYFGRKLAVDASSTYFLLVCGTCKEHTITNAGE 60
 DB 1 MGKGLTSOVIADNAKSAIRKVENKAFGRVIAVDSKCTYOFLLAVRQDSQ-LGSEEGE 59
 QY 51 VTSHQGPNKRTIRLEAGIKPVYVFGKPPDMKROGCAKRYSKRDYATKLLIYAVVGD 120
 DB 60 TISILMKIKRTYRKFENGVKVFYVDFUKYDPMKGGLERKSERRAEKALIFAKKRD 119

```

QY 121 KDAIEKLSKRTVAVTROHNEDECKRLRLKMGVPVVAPEAFACALCINCKVFAVASED 180
DB 118 EEEVFKETKRIYVVTQIHNDCHCLHSLKQIPYLAPEASBASCALAKACVVAASND 177
QY 181 MSLLTGAPFRLHLDPSKRIIPVEFPYAVLEELTMDQFIDLCGLCCDYCDSTK 240
DB 178 MCLTFSPLMLHSLASEAKRIPTGFHLSRYLQFELGNQFQFDCLDLSDYCESIA 237
QY 241 GIGGOTALKLRHGSISLELNLKNDROYQPEDMYOCARLFFEPNVT-LDIFLKMT 299
DB 238 GIGAKRAVDLIQHKSLIEELVRLOPSKIPVPEWMLHKEAQJLFEPEVQESVELKWS 297
QY 300 APDEGLISFLVNDGNFEDRVKALEIKISAKNKSOGLESPFKPATTISAPLSKFT 359
DB 298 EPNEEELVKMCEKQFSEERISISGVKRLSKNSQSTQGRUDLPFNVCSSLS-AKREKP 356
QY 360 SDKSKRAANKTKKAGKK 379
DB 357 E---PKGPARKKAKTGAGK 373

RESULT 14
C96830
ID 096830 PRELIMINARY: PRT: 385 AA.
AC 096830:
DT 01-MAY-1999 (TEMBLrel: 10, Created)
DT 01-MAY-1999 (TEMBLrel: 10, last sequence update)
DT 01-OCT-2002 (TEMBLrel: 22, last annotation update)
DE EG:EG0003.3 protein.
GN FEH1 OR EG:EG0003.3 OR CG8648.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Spizididae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY:
RX MEDLINE=20196006; PubMed=10731132:
RA Adams M.E., Gelikier S.E., Holt R.A., Evans C.A., Gacyayne J.D.,
RA Anagnostides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle S.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazes R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter P.S., Holt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews Plankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram R.P., Bhattacharya D., Bolshakov S.,
RA Botkova D., Botchan M.R., Bouck J., Brokslein P., Brothier P.,
RA Britis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra S.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos R., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Danko B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischman W.,
RA Fodler C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Gloor A., Gong F., Gorell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman J.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howard T.J., Wei M.-H., Ibeagwa C.,
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Metcalfe G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA Palazuelo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.S.,
RA Reinert K., Remington K., Saunders R.D.C., Scheefer F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M.C., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

```

```

RA Ye J., Yeh R.-F., Zaveri J.S., Zhu M., Zhang S., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.X., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.K., Venter J.C.:
RA "The genome sequence of Drosophila melanogaster."
RA Science 287:2185-2195(2000).
RA (2)
RP SEQUENCE FROM N.A.
RA Murphy L., Harris D., Barrett J.:
RA "Sequencing the distal X chromosome of Drosophila melanogaster."
RA Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.
RN 131
RP SEQUENCE FROM N.A.
RA bios P.:
RA Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DB EMBL: AE003805; AF57544.1;
DB EMBL: AF031863; CA21120.1;
DB HSPB: Q58839; 1A76.
DB FLYBase: FBgn0025852; FEH1.
DB InterPro: IPR002421; 5_3_exonuclease.
DB InterPro: IPR00513; EXO_N_1.
DB InterPro: IPR003583; HHH_1.
DB InterPro: IPR003584; HHH_2.
DB InterPro: IPR006084; XPGC_Rad.
DB InterPro: IPR006085; XPG_L.
DB InterPro: IPR006086; XPG_N.
DB Pfam: PF01367; 5_3_exonuclease; 1.
DB Pfam: PF00867; XPG_L; 1.
DB Pfam: PF00752; XPG_N; 1.
DB PRINTS: PR0053; XPGRADSUPER.
DB SMART: SM00276; HNH1; 1.
DB SMART: SM00279; HNH2; 1.
DB SMART: SM00484; XPG1; 1.
DB SMART: SM00485; XPGN; 1.
DB PROSITE: PS0084; XPG_1; 1.
DB PROSITE: PS00842; XPG_2; 1.
SQ SEQUENCE 385 AA; 42548 MW; DDA0831C2BDA9240 CRC64;

Query Match 5.5%; Score 998.5; DR 5; Length 395;
Best Local Similarity 52.9%; Pctd. No. 5.2e-65;
Matches 202; Conservative 65; Mismatches 104; Indels 11; Gaps

QY 1 MGICQJTKLIADNAPPAKMKPFKPFCKKIVNDASKSTYQFLIVSGTGMETTCNFAE 6
DB 1 MGICQJTKLIADNAPPAKMKPFKPFCKKIVNDASKSTYQFLIVSGTGMETTCNFAE 6
QY 61 VSHLQGMSEKTRIL:EAQIKPYPYVQAKPPMKKELAKKSKRQDA-KULTEAVFVC 1
DB 61 VSHLQGMSEKTRIL:EAQIKPYPYVQAKPPMKKELAKKSKRQDA-KULTEAVFVC 1
DB 60 PSHLMQMEFRTIK:LTNGIKFPYVVDGKPPDKSGELAKFAFREAEKALKAAITDAGD 1
QY 121 KDAIEKLSKRTVAVTROHNEDECKRLRLKMGVPVVAPEAFACALCINCKVFAVASED 1
DB 121 KDAIEKLSKRTVAVTROHNEDECKRLRLKMGVPVVAPEAFACALCINCKVFAVASED 1
QY 181 MSLLTGAPFRLHLDPSKRIIPVEFPYAVLEELTMDQFIDLCGLCCDYCDSTK 2
DB 181 MSLLTGAPFRLHLDPSKRIIPVEFPYAVLEELTMDQFIDLCGLCCDYCDSTK 2
QY 241 GIGGOTALKLRHGSISLELNLKNDROYQPEDMYOCARLFFEPNVT-LDIFLKMT 2
DB 241 GIGGOTALKLRHGSISLELNLKNDROYQPEDMYOCARLFFEPNVT-LDIFLKMT 2
QY 238 GIGAKRAVDLIQHKSLIEELVRLOPSKIPVPEWMLHKEAQJLFEPEVQESVELKWS 2
DB 238 GIGAKRAVDLIQHKSLIEELVRLOPSKIPVPEWMLHKEAQJLFEPEVQESVELKWS 2
QY 300 APDEGLISFLVNDGNFEDRVKALEIKISAKNKSOGLESPFKPATTISAPLSKFT 3
DB 300 APDEGLISFLVNDGNFEDRVKALEIKISAKNKSOGLESPFKPATTISAPLSKFT 3
QY 298 EPNEEELVKMCEKQFSEERISISGVKRLSKNSQSTQGRUDLPFNVCSSLS-AKREKP 3
DB 298 EPNEEELVKMCEKQFSEERISISGVKRLSKNSQSTQGRUDLPFNVCSSLS-AKREKP 3
QY 360 SDKSKRAANKTKKAGKK 376
DB 360 SDKSKRAANKTKKAGKK 376
QY 357 E---PKGPARKKAKTGAGK 373
DB 357 E---PKGPARKKAKTGAGK 373

RESULT 15
C96830
ID 096830 PRELIMINARY: PRT: 650 AA.

```

AC Q96Z01; 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DI 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Flap endonuclease-1.
GN FEN1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Casta J.J. III, Schmutte C., Taraschi T.F.;
RT *Flap Endonuclease-1 (Plasmodium falciparum).
PL Submitted (JUN-2000) to the EMBL/GenBank/CCDS databases.
DR EMBL: AC278264; AAC01445.1; .
DR HSSP: Q58839; A76.
DR InterPro: IPR002421; 5_3-exonuclease.
DR InterPro: IPR000513; Exo_N_F.
DR InterPro: IPR003584; HNH_2.
DR InterPro: IPR006084; XPG_Kud.
DR InterPro: IPR006086; XPG_I.
DR InterPro: IPR006085; XPG_N.
DR Pfam: PF01567; 5_3-exonuclease; 1.
DR Pfam: PF00867; XPG_1; .
DR Pfam: PF00752; XPG_N; 1.
DR PRINTS: PR00853; XPGRADSUPER.
DR SMART: SM00279; HNH2; .
DR SMART: SM00484; XPG1; .
DR SMART: SM00485; XPGN; 1.
DR Endonuclease.
KW SEQUENCE 650 AA; 73989 MW; 2F9464DAB7F45B1D CRC64;

Query Match 45.7%; Score 887; DB 5; Length 650;
Best Local Similarity 46.6%; Pred. No. 1.5e-56;
Matches 164; Conservative 74; Mismatches 117; Indels 20; Gaps 5;

QY 1 MGKIGTKLLADAPNAKMEQKFESEYFGKIAVDASMSYYOFLIVGRGM-EILTNEAG 59
DB 1 MGKIGTKTADAPNAIKIEISLMGIINIDASMSLYQFTINRDEQGNCTNPSG 60

QY 60 EYSHQGMFNTIRLLENGIKPVYVFCGKPPDMKKQELAKRYSKRDATKDL-EAVVQ 119
DB 61 EYTSH-SGLMSRSITRMENGLKPIYVFDGAPPELKGSELEKRGKRAAEIIRAKKEEG 120

QY 120 DKDALEKLSKRTVYKTRQHNEDCKRLRLMGVYVEAPSEAEACALCNDKVPVASE 175
DB 121 NLEIKKQSGRTVYTRKQNEEAKKLLTLMGIPITFAPCEAECSQAFILRYLAATATE 180

QY 180 DMDSLTFCAPFLRLHL-----MDPSKKITPVMEFDYAKVLEELTMDGFIDLC 228
DB 181 DADALVEFGTKILIRNLNANATSNQNNKNSKRGYILTEINDEYIKGINLIMDETFDC 240

QY 229 ILGGCDYCSIKIGGOTALK-LRQGSIESILENLNKRQCIPEDMVYQEARRLFKEPN 288
DB 241 ILGGCDYCDITKIGSKTAYNLIKENCTEKTENIDOKYOVPSNRPQEKRSFIPDN 300

QY 289 VTIIDIP---ELKWTAPDEGLISFLVNDGCFNEDRVTKALEIKSANKKSSQGLSEEF 344
DB 301 V---LPKEDIKTIDWNEPQIEELKHFLIKYNNELRVNTYINRLKARKVT-QRRIDNPF 357

QY 345 KPIATTSAPLARKET-SDKTSKAANKKTKAGGK 378
DB 358 TACTKKSTKLIVEETKKEOTLPAKGGKRRPTAGDK 392

Search completed: October 22, 2003, 12:42:37
Job time : 104 secs

Db 361 DKTSKAANKTKAGCKK 379

RESULT 2

US-09-426-557-6
 : Sequence 6, Application US/09426557
 : Patent No. 6232527
 : GENERAL INFORMATION:
 : APPLICANT: Mahajan, Pramod B.
 : TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
 : TITLE OF INVENTION: thereof
 : FILE REFERENCE: 0961
 : CURRENT APPLICATION NUMBER: US/09/426,557
 : CURRENT FILING DATE: 1999-10-22
 : EARLIER APPLICATION NUMBER: 60/112,332
 : EARLIER FILING DATE: 1998-12-15
 : NUMBER OF SEQ ID NOS: 10
 : SOFTWARE: FastSeq for Windows Version 3.0
 : SEQ ID NO: 6
 : LENGTH: 379
 : TYPE: PRI
 : ORGANISM: Zea mays
 : US-09-426-557-6

Query Match: 100.0%; Score 1939; DB 3; Length 379;
 Best Local Similarity 100.0%; Pred. No. 1,9e+183;
 Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGICGLTKLADNAPKAKKEQFESEYFGKRIADVDSMSIYQFLIVGRGCMETL;NFACE 60
 |||||||
 DB 1 MGICGLTKLADNAPKAKKEQFESEYFGKRIADVDSMSIYQFLIVGRGCMETL;NFACE 60
 |||||||
 QY 61 VTSHLOGMFRITRLLEAGIKIPYVVDGKPPDMKKOELAKRYSKRDATKDLTEAVEVGD 123
 |||||||
 DB 61 VTSHLOGMFRITRLLEAGIKIPYVVDGKPPDMKKOELAKRYSKRDATKDLTEAVEVGD 123
 |||||||
 QY 121 KDAIEKLSKRTVAVTQHNEDECKRLRLMGVYVVEAPSEAEACALCINDKVFVAVASD 189
 |||||||
 DB 121 KDAIEKLSKRTVAVTQHNEDECKRLRLMGVYVVEAPSEAEACALCINDKVFVAVASD 189
 |||||||
 QY 181 MDLTFGAPRFLRLHMDPSSKRIYVMEFDVAKYLEBELTMDQFIDLCICGDCYDSIK 240
 |||||||
 DB 181 MDLTFGAPRFLRLHMDPSSKRIYVMEFDVAKYLEBELTMDQFIDLCICGDCYDSIK 240
 |||||||
 QY 241 GIGGOTALKLIRHGSTESTILENLNCDROYIPEDMWYQEARRLFKEPNVTLIDIPELKMTA 306
 |||||||
 DB 241 GIGGOTALKLIRHGSTESTILENLNCDROYIPEDMWYQEARRLFKEPNVTLIDIPELKMTA 306
 |||||||
 QY 301 PDEGLISFLVKDNGFNEEDRVTALEKIKSAKSSQGRLESFFKPTATTSAPLKRKETS 360
 |||||||
 DB 301 PDEGLISFLVKDNGFNEEDRVTALEKIKSAKSSQGRLESFFKPTATTSAPLKRKETS 360
 |||||||
 QY 361 DKTSKAANKTKAGCKK 379
 |||||||
 DB 361 DKTSKAANKTKAGCKK 379
 |||||||

RESULT 3

US-09-426-557-4
 : Sequence 4, Application US/09426557
 : Patent No. 6232527
 : GENERAL INFORMATION:
 : APPLICANT: Mahajan, Pramod B.
 : TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
 : TITLE OF INVENTION: thereof
 : FILE REFERENCE: 0961
 : CURRENT APPLICATION NUMBER: US/09/426,557
 : CURRENT FILING DATE: 1999-10-22
 : EARLIER APPLICATION NUMBER: 60/112,332
 : EARLIER FILING DATE: 1998-12-15
 : NUMBER OF SEQ ID NOS: 10
 : SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 4
 : LENGTH: 379
 : TYPE: PRI
 : ORGANISM: Zea mays
 : US-09-426-557-4

Query Match: 99.7%; Score 1933; DB 3; Length 379;
 Best Local Similarity 99.7%; Pred. No. 7.6e+183;
 Matches 379; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGICGLTKLADNAPKAKKEQFESEYFGKRIADVDSMSIYQFLIVGRGCMETL;NFACE 60
 |||||||
 DB 1 MGICGLTKLADNAPKAKKEQFESEYFGKRIADVDSMSIYQFLIVGRGCMETL;NFACE 60
 |||||||
 QY 61 VTSHLOGMFRITRLLEAGIKIPYVVDGKPPDMKKOELAKRYSKRDATKDLTEAVEVGD 123
 |||||||
 DB 61 VTSHLOGMFRITRLLEAGIKIPYVVDGKPPDMKKOELAKRYSKRDATKDLTEAVEVGD 123
 |||||||
 QY 121 KDAIEKLSKRTVAVTQHNEDECKRLRLMGVYVVEAPSEAEACALCINDKVFVAVASD 189
 |||||||
 DB 121 KDAIEKLSKRTVAVTQHNEDECKRLRLMGVYVVEAPSEAEACALCINDKVFVAVASD 189
 |||||||
 QY 181 MDLTFGAPRFLRLHMDPSSKRIYVMEFDVAKYLEBELTMDQFIDLCICGDCYDSIK 240
 |||||||
 DB 181 MDLTFGAPRFLRLHMDPSSKRIYVMEFDVAKYLEBELTMDQFIDLCICGDCYDSIK 240
 |||||||
 QY 241 GIGGOTALKLIRHGSTESTILENLNCDROYIPEDMWYQEARRLFKEPNVTLIDIPELKMTA 306
 |||||||
 DB 241 GIGGOTALKLIRHGSTESTILENLNCDROYIPEDMWYQEARRLFKEPNVTLIDIPELKMTA 306
 |||||||
 QY 301 PDEGLISFLVKDNGFNEEDRVTALEKIKSAKSSQGRLESFFKPTATTSAPLKRKETS 360
 |||||||
 DB 301 PDEGLISFLVKDNGFNEEDRVTALEKIKSAKSSQGRLESFFKPTATTSAPLKRKETS 360
 |||||||
 QY 361 DKTSKAANKTKAGCKK 379
 |||||||
 DB 361 DKTSKAANKTKAGCKK 379
 |||||||

RESULT 4

US-09-426-557-6
 : Sequence 8, Application US/09426557
 : Patent No. 6232527
 : GENERAL INFORMATION:
 : APPLICANT: Mahajan, Pramod B.
 : TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
 : TITLE OF INVENTION: thereof
 : FILE REFERENCE: 0961
 : CURRENT APPLICATION NUMBER: US/09/426,557
 : CURRENT FILING DATE: 1999-10-22
 : EARLIER APPLICATION NUMBER: 60/112,332
 : EARLIER FILING DATE: 1998-12-15
 : NUMBER OF SEQ ID NOS: 10
 : SOFTWARE: FastSeq for Windows Version 3.0
 : SEQ ID NO: 8
 : LENGTH: 379
 : TYPE: PRI
 : ORGANISM: Zea mays
 : US-09-426-557-6

Query Match: 99.7%; Score 1933; DB 3; Length 379;
 Best Local Similarity 99.7%; Pred. No. 7.6e+183;
 Matches 379; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGICGLTKLADNAPKAKKEQFESEYFGKRIADVDSMSIYQFLIVGRGCMETL;NFACE 60
 |||||||
 DB 1 MGICGLTKLADNAPKAKKEQFESEYFGKRIADVDSMSIYQFLIVGRGCMETL;NFACE 60
 |||||||
 QY 61 VTSHLOGMFRITRLLEAGIKIPYVVDGKPPDMKKOELAKRYSKRDATKDLTEAVEVGD 123
 |||||||
 DB 61 VTSHLOGMFRITRLLEAGIKIPYVVDGKPPDMKKOELAKRYSKRDATKDLTEAVEVGD 123
 |||||||
 QY 121 KDAIEKLSKRTVAVTQHNEDECKRLRLMGVYVVEAPSEAEACALCINDKVFVAVASD 189
 |||||||
 DB 121 KDAIEKLSKRTVAVTQHNEDECKRLRLMGVYVVEAPSEAEACALCINDKVFVAVASD 189
 |||||||


```

Db      221 KDAIEKLSKRTVKTROHNEDECKRLRLMGVYVVEAPSEAECAALCTINDKFAVASED 180
Cy      181 MSTLTFGAPRFLRLHMDPSSKRIPIWEPDVAKVLELELTDQFIDCITLGGCDYCSIK 240
Db      181 KDSITFGAPRFLRLHMDPSSKRIPIWEPDVAKVLELELTDQFIDCITLGGCDYCSIK 240
Cy      241 GIGGOTALKLIRHOSISISILENLKDRQYQPEDKPYOEARLEFEKPVNTLIDIFLXKTA 300
Db      241 GIGGOTALKLIRHOSISISILENLKDRQYQPEDKPYOEARLEFEKPVNTLIDIFLXKTA 300
Cy      301 PDEGLISFVKNQNGFNEDRVKALEKISAKNKSOGRLSEFFKPTATYSAPJAKRKETS 360
Db      301 PDEGLISFVKNQNGFNEDRVKALEKISAKNKSOGRLSEFFKPTATYSAPJAKRKETS 360
Cy      361 DKTSKAANKTKKAGKKK 379
Db      361 DKTSKAANKTKKAGKKK 379

```

```

RESULT 5
US-08-823-516-137
Sequence 137, Application US/08823516
Patent No. 5994069
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
APPLICANT: Lyamichev, Victor I.
APPLICANT: Most, Andrea L.
APPLICANT: Brown, Mary Ann D.
TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medien 6 Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/756,386
FILING DATE: 28-NOV-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/559,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Lyollis, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: F0RS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
FAX: (415) 397-8338

```

```

INFORMATION FOR SEQ ID NO: 137:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-08-823-516-137

```

```

Query Match: 53.2%, Score 1032, DB 2, Length 380;
Best Local Similarity 53.8%, Prod. No. 1e-93;
Matches 206; Conservative 69; Mismatches 98; Indels 10; Gaps

```

```

Cy      1 MGIKGLKILADNAPKAKKQGEKFSYGRKILAVASMSIVCFLLVSGTGMETLINNGE
Db      1 MGIKGLKILADNAPKAKKQGEKFSYGRKILAVASMSIVCFLLVSGTGMETLINNGE
Cy      61 VTSIIQGFNFTIRLENGIKPVYVVDKPPKPKKQKFLAKRYSKRDATKIDTAVEYGD
Db      60 TTSILKMEYKTIIMMENGKIFVYVVDKPPKPKKQKFLAKRYSKRDATKIDTAVEYGD
Cy      121 KDAIEKLSKRTVKTROHNEDECKRLRLMGVYVVEAPSEAECAALCTINDKFAVASED
Db      120 EDEVEKELTKRLVAVKQHNQCKHILSLMGVYVVEAPSEAECAALCTINDKFAVASED
Cy      181 MSTLTFGAPRFLRLHMDPSSKRIPIWEPDVAKVLELELTDQFIDCITLGGCDYCSIK
Db      180 MDSLTFGAPRFLRLHMDPSSKRIPIWEPDVAKVLELELTDQFIDCITLGGCDYCSIK
Cy      241 GIGGOTALKLIRHOSISISILENLKDRQYQPEDKPYOEARLEFEKPVNTLIDIFLXKTA
Db      240 GIGGOTALKLIRHOSISISILENLKDRQYQPEDKPYOEARLEFEKPVNTLIDIFLXKTA
Cy      299 TAPDEGLISFVKNQNGFNEDRVKALEKISAKNKSOGRLSEFFKPTATYSAPJAKRKETS
Db      299 SEPNEELIKRFGCKGKPSDEERPSGVKRLSKSOGSTGKADGPFKVTGLSS-ARRKE
Cy      359 TSCK--TSKAANKTKKAGKKK 379
Db      358 PEKSKTKKKA--KTGAKGKFK 377

```

```

RESULT 6
US-09-426-557-10
Sequence 10, Application US/09426557
Patent No. 5232527
GENERAL INFORMATION:
APPLICANT: Mahajan, Pramod B.
TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
FILE REFERENCE: 0961
CURRENT APPLICATION NUMBER: US/09/426,557
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 60/112,332
EARLIER FILING DATE: 1998-12-15
SOFTWARE: FASTSTRO for Windows Version 3.0
SEQ ID NO: 10
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-09-426-557-10

```

```

Query Match: 53.2%, Score 1032, DB 3, Length 380;
Best Local Similarity 53.8%, Prod. No. 1e-93;
Matches 206; Conservative 69; Mismatches 98; Indels 10; Gaps

```

```

Cy      1 MGIKGLKILADNAPKAKKQGEKFSYGRKILAVASMSIVCFLLVSGTGMETLINNGE
Db      1 MGIKGLKILADNAPKAKKQGEKFSYGRKILAVASMSIVCFLLVSGTGMETLINNGE
Cy      61 VTSIIQGFNFTIRLENGIKPVYVVDKPPKPKKQKFLAKRYSKRDATKIDTAVEYGD
Db      60 TTSILKMEYKTIIMMENGKIFVYVVDKPPKPKKQKFLAKRYSKRDATKIDTAVEYGD
Cy      121 KDAIEKLSKRTVKTROHNEDECKRLRLMGVYVVEAPSEAECAALCTINDKFAVASED
Db      120 EDEVEKELTKRLVAVKQHNQCKHILSLMGVYVVEAPSEAECAALCTINDKFAVASED
Cy      181 MSTLTFGAPRFLRLHMDPSSKRIPIWEPDVAKVLELELTDQFIDCITLGGCDYCSIK
Db      180 MDSLTFGAPRFLRLHMDPSSKRIPIWEPDVAKVLELELTDQFIDCITLGGCDYCSIK
Cy      241 GIGGOTALKLIRHOSISISILENLKDRQYQPEDKPYOEARLEFEKPVNTLIDIFLXKTA
Db      240 GIGGOTALKLIRHOSISISILENLKDRQYQPEDKPYOEARLEFEKPVNTLIDIFLXKTA
Cy      299 TAPDEGLISFVKNQNGFNEDRVKALEKISAKNKSOGRLSEFFKPTATYSAPJAKRKETS
Db      299 SEPNEELIKRFGCKGKPSDEERPSGVKRLSKSOGSTGKADGPFKVTGLSS-ARRKE
Cy      359 TSCK--TSKAANKTKKAGKKK 379
Db      358 PEKSKTKKKA--KTGAKGKFK 377

```

Db 60 TTSHLMGMFRTIRMENGICRPVYFDCKPPQLASGLAKSERRAEKQLQQAAGA 119
 QY 121 KDAIFKLSKRTVYVROHNEDCKRLRLMGVYVVEAPSEAEACALCTNDKVFVASED 180
 Db 120 EGVVEKFTKRLVYVROHNEDCKRLRLMGVYVVEAPSEAEACALCTNDKVFVASED 179
 QY 181 MDLTFGAPFLRLHMLDPSKKIPVMEFVAKVLEELFLTMDQFIDCTICGCGYCDSTK 240
 Db 180 MDLTFGAPFLRLHMLDPSKKIPVMEFVAKVLEELFLTMDQFIDCTICGCGYCDSTK 239
 QY 241 GIGSOTALKLIRHOSIESILENLKDKRYQIPEDMPYQEARLKEPNTLD--1PELKW 298
 Db 240 GIGSOTALKLIRHOSIESILENLKDKRYQIPEDMPYQEARLKEPNTLD--1PELKW 298
 QY 299 TAPDEGLIFLVKNGFNEDVTKAIEKISAKNNSSQGRLESEFKPTATTSAPLKRKE 358
 Db 299 SEPNEEDLIFMGCGKQFSEERIRSGVRLSKRSGSTOGRDLDFFKVTGSLSS-AKRKE 357
 QY 359 TSDK--TSKAANKTKAGKKK 379
 Db 358 PEKGGSTKKA---KTGAAGKK 377

RESULT 7
 US-08-455-968E-1
 Sequence 1, Application US/08455968E

Patent No. 5874283
 GENERAL INFORMATION:
 APPLICANT: Harrington, John L.
 APPLICANT: Hsieh, Chih-Lin
 APPLICANT: Lieber, Michael
 TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
 NUMBER OF SEQUENCES: 63
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend and Crew LLP
 STREET: Two Embarcadero Center, 9th Floor
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94111-3834
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/455,968E
 FILING DATE: 30-MAY-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Apple, Randolph T.
 REGISTRATION NUMBER: 36,429
 REFERENCE/DOCKET NUMBER: 18995-000100
 TELEPHONE: 415-576-0200
 TELEFAX: 415-576-0300
 INFORMATION FOR SEQ. ID NO. 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 380 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-455-968E-1

Query Match 53.2% Score 1031; DB 2; Length 380;
 Best Local Similarity 53.8% Pred. No. 1,3e-93;
 Matches 206; Conservative 69; Mismatches 98; Indels 10; Gaps 6;
 QY 1 MGIGLTLKACNAKKEKFESEYGRKTAAYDQMSISYQFLIYVGTGSETLTNAKE 60
 Db 1 MGIGLTLKADVAAPSAIRENDIKSTPGRKVAIDMSISYQFLIYV-ROGSDVIONERGE 59

QY 61 VTSHLMGMENRTIRLEAGIKRPVYVEDGKPPDKKOELAKRYSKRDATKDLTEAVEVGD 119
 Db 60 TTSHLMGMFRTIRMENGICRPVYFDCKPPQLASGLAKSERRAEKQLQQAAGA 119
 QY 121 KDAIFKLSKRTVYVROHNEDCKRLRLMGVYVVEAPSEAEACALCTNDKVFVASED 180
 Db 120 EGVVEKFTKRLVYVROHNEDCKRLRLMGVYVVEAPSEAEACALCTNDKVFVASED 179
 QY 181 MDLTFGAPFLRLHMLDPSKKIPVMEFVAKVLEELFLTMDQFIDCTICGCGYCDSTK 240
 Db 180 MDLTFGAPFLRLHMLDPSKKIPVMEFVAKVLEELFLTMDQFIDCTICGCGYCDSTK 239
 QY 241 GIGSOTALKLIRHOSIESILENLKDKRYQIPEDMPYQEARLKEPNTLD--1PELKW 298
 Db 240 GIGSOTALKLIRHOSIESILENLKDKRYQIPEDMPYQEARLKEPNTLD--1PELKW 298
 QY 299 TAPDEGLIFLVKNGFNEDVTKAIEKISAKNNSSQGRLESEFKPTATTSAPLKRKE 358
 Db 299 SEPNEEDLIFMGCGKQFSEERIRSGVRLSKRSGSTOGRDLDFFKVTGSLSS-AKRKE 357
 QY 359 TSDK--TSKAANKTKAGKKK 379
 Db 358 PEKGGSTKKA---KTGAAGKK 377

RESULT 8
 US-08-823-516-138
 Sequence 138, Application US/08823516
 Patent No. 5994053

GENERAL INFORMATION:
 APPLICANT: Hall, Jeff G.
 APPLICANT: Lyatchev, Victor I.
 APPLICANT: Mast, Andrea L.
 APPLICANT: Brow, Mary Ann D.
 TITLE OF INVENTION: Detection of Nucleic Acids By Multiple
 NUMBER OF SEQUENCES: 163
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Moden & Carroll, LLP
 STREET: 220 Montgomery Street, Suite 2200
 CITY: San Francisco
 STATE: California
 COUNTRY: United States of America
 ZIP: 94104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/823,516
 FILING DATE: 24-MAR-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US97/01072
 FILING DATE: 21-JAN-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/759,078
 FILING DATE: 02-DEC-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/758,314
 FILING DATE: 02-DEC-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/756,365
 FILING DATE: 29-NOV-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/682,653
 FILING DATE: 12-JUL-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/599,493
 FILING DATE: 24-JAN-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Ingolia, Diane E.

REGISTRATION NUMBER: 40,027
 REFERENCE/DOCKET NUMBER: FORS-02736
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 705-8410
 TELEFAX: (415) 397-8338
 INFORMATION FOR SEQ ID NO: 138:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 378 amino acids
 TYPE: amino acid
 STRANDEDNESS: not relevant
 TOPOLOGY: not relevant
 MOLECULE TYPE: protein
 US-09-823-516-138

Query Match

52.38: Score 1014; DB 2: Length 378;
 Best Local Similarity 53.24; Pred. No. 6, le-92;
 Matches 202: Conservative 69; Mismatches 101; Indels 5; Gaps 6;

QY 1 MGIRGLTKLLADNAPKAMKQKFESEYFGKRIADASKSIYQFLIVGRTGMEITINEAGE 40
 DB 1 MGIRGLTKLLADNAPKAMKQKFESEYFGKRIADASKSIYQFLIVGRTGMEITINEAGE 59
 QY 61 VISHLOGMFNRTIRLEAGIKPYVVDGKPPDKKQKQELAKRYSKRDQATKDYTEAVEYGD 120
 DB 60 TTS-LMGFPRITR-MENGIRKPYVVDGKPPDKKQKQELAKRYSKRDQATKDYTEAVEYGD 117
 QY 121 KDAIEKSKRCVAVTRQHNEDCKRLRLMGVPPVEAPSEAEACALCINDKVFVASED 180
 DB 118 BEVEKFTKRLVYVTKQHNEDCKRLRLMGVPPVEAPSEAEACALCINDKVFVASED 177
 QY 181 MDLTFGAPRFLRLDPPSSKIPVMEFDVAKYLELEITMDQFLDCLTCGQDYDQSDIR 240
 DB 178 MDCLTFGAPRFLRLDPPSSKIPVMEFDVAKYLELEITMDQFLDCLTCGQDYDQSDIR 237
 QY 241 GIGGQALKLIRQGSIESTLENLNDKRDYIPEDMPYQEARRLPKKPNVT-IDIPELKMT 299
 DB 238 GIGAKRAVDLQKHSLEIVRLDPSKYVPVENMLHKEAQQLFLPEPVYDSEVELKMS 297
 QY 300 APDEGLISFLVKNQNGENRQVTKALEIKSAKNSQGLSEFEKPYATSAPLKPKET 359
 DB 298 EPNEELVKFMCGEKQSEERIRSGVRLSKSQSGTQGRLDQFVKVTSLS-ARKKRP 356
 QY 360 SDKTSKAANKTKAGCKK 379
 DB 357 E--PKGPARKKAKTGAGK 373

RESULT 9

US-08-455-968E-3

Sequence 3, Application US/08455968E

Patent No. 5874283

GENERAL INFORMATION:

APPLICANT: Harrington, John L.

APPLICANT: Hsieh, Chih-Lin

APPLICANT: Lieber, Michael

TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease

NUMBER OF SEQUENCES: 63

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/455,968E

FILING DATE: 30-MAY-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
 NAME: Apple, Randolph T.
 REGISTRATION NUMBER: 36,429
 REFERENCE/DOCKET NUMBER: 18985-000100
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-576-0200
 TELEFAX: 415-576-0300
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 377 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-455-968E-3

Query Match: 52.08: Score 1007.5; DB 2: Length 377;
 Best Local Similarity 53.44; Pred. No. 2, 7e-91;
 Matches 203: Conservative 68; Mismatches 100; Indels 5; Gaps 6;

QY 1 MGIRGLTKLLADNAPKAMKQKFESEYFGKRIADASKSIYQFLIVGRTGMEITINEAGE 40
 DB 1 MGIRGLTKLLADNAPKAMKQKFESEYFGKRIADASKSIYQFLIVGRTGMEITINEAGE 59
 QY 61 VISHLOGMFNRTIRLEAGIKPYVVDGKPPDKKQKQELAKRYSKRDQATKDYTEAVEYGD 120
 DB 60 TTS-LMGFPRITR-MENGIRKPYVVDGKPPDKKQKQELAKRYSKRDQATKDYTEAVEYGD 117
 QY 121 KDAIEKSKRCVAVTRQHNEDCKRLRLMGVPPVEAPSEAEACALCINDKVFVASED 180
 DB 117 MEVEKFTKRLVYVTKQHNEDCKRLRLMGVPPVEAPSEAEACALCINDKVFVASED 177
 QY 181 MDLTFGAPRFLRLDPPSSKIPVMEFDVAKYLELEITMDQFLDCLTCGQDYDQSDIR 240
 DB 178 MDCLTFGAPRFLRLDPPSSKIPVMEFDVAKYLELEITMDQFLDCLTCGQDYDQSDIR 237
 QY 241 GIGGQALKLIRQGSIESTLENLNDKRDYIPEDMPYQEARRLPKKPNVT-IDIPELKMT 299
 DB 237 GIGAKRAVDLQKHSLEIVRLDPSKYVPVENMLHKEAQQLFLPEPVYDSEVELKMS 297
 QY 300 APDEGLISFLVKNQNGENRQVTKALEIKSAKNSQGLSEFEKPYATSAPLKPKET 359
 DB 297 EPNEELVKFMCGEKQSEERIRSGVRLSKSQSGTQGRLDQFVKVTSLS-ARKKRP 356
 QY 360 SDKTSKAANKTKAGCKK 379
 DB 356 EPKGS--AKKAKTQAGK 372

RESULT 10

US-08-455-968E-3

Sequence 10, Application US/08455968E

Patent No. 5874283

GENERAL INFORMATION:

APPLICANT: Harrington, John L.

APPLICANT: Hsieh, Chih-Lin

TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease

NUMBER OF SEQUENCES: 63

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/455,968E

FILING DATE: 30-MAY-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Apple, Randolph T.
 REGISTRATION NUMBER: 36,429
 REFERENCE/DOCKET NUMBER: 18985-000100
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-576-0200
 TELEFAX: 415-576-0300
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 378 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-455-968E-10

Query Match 51.4%; Score 996; DB 2; Length 378;
 Best Local Similarity 52.4%; Pred. No. 3,7e+90;
 Matches 199; Conservative 70; Mismatches 103; Indels 8; Gaps 6;

QY 1 MGIGLTKLADNAPKAKKEQKESYGRKAVDASMSIYQFLIVGRGTGMLTINAGE 60
 DB 1 MEHGIKALADVAPSAIRENDIKSYGRKVAIDASMSIYQFLIAY-RQGADVLQNEGE 59
 QY 61 VTSHLQGMENFTIRLENGIKPVYVFDGKPPDMKKQELAKRYSKRDATKDCLEAVEVG 120
 DB 60 TTS-LMGMYRTMR-MENGIKPVYVFDGKPPQLKSGELAKSERAEKQJQQAQAGM 117
 QY 121 KDAIEKLSKRTVKTROHNEDECKRLRLMSVPVVEAPSEAECAALCINDKVAVASED 180
 DB 118 EEVEKEFTRLVKTAKYKHNDCKHLLSLMGIPYLDAPSEASQCALAKKGVAAATFD 177
 QY 181 MSLTEGAPRFLRLHMDPSSKKIPVMEFDVAKVLEELTMDQFIDLCILGCGCYCSIK 240
 DB 178 MDCLIFGSPVAKRHLIASEAKKLIQEFHLSRVLOELQNEQVDCILGSPCYCSIR 237
 QY 241 GIGGICLALKLROHGSLESLINKDKYQIPEDMFTYQEARCKEPPNVT-LDIPELKWT 299
 DB 238 GIGAKRAVDLIQKHISIEIVRRGDPKSYVPEXWLMHEAQCLPELEVVDPSEVFLKMS 297
 QY 300 APDEGLSFLVKDNGENEDRVTAIEIKSAKAKSSQGRLESPFKPATISALAKREE 359
 DB 298 EPNEBELVKTMGEGEQFEERIRISGVKRLSKRSGSTOGRLDDEFFKV7GSLSS-AKKEEP 356
 QY 360 SDKTSKAAANKKTKAGCKKK 379
 DB 357 F---PKGPAAKKKAKTGAAGK 373

RESULT 11

US-08-455-968E-5
 Sequence 5, Application US/08455968E
 Patent No. 5874283
 GENERAL INFORMATION:
 APPLICANT: Harrington, John L.
 APPLICANT: Hsieh, Michael
 APPLICANT: Lieder, Michael
 TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
 NUMBER OF SEQUENCES: 63
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend and Crew LLP
 STREET: Two Embarcadero Center, 8th floor
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94111-3834
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: 09/08/455,968E
 FILING DATE: 30-MAY-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Apple, Randolph T.
 REGISTRATION NUMBER: 36,429
 REFERENCE/DOCKET NUMBER: 18985-000100
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-576-0200
 TELEFAX: 415-576-0300
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 382 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-455-968E-5

Query Match 48.4%; Score 938; DB 2; Length 382;
 Best Local Similarity 49.5%; Pred. No. 2e+84;
 Matches 190; Conservative 65; Mismatches 109; Indels 16; Gaps

QY 1 MGIGLTKLADNAPKAKKEQKESYGRKAVDASMSIYQFLIVGRGTGMLTINAGE 60
 DB 1 MEHGIKALADVAPSAIRENDIKSYGRKVAIDASMSIYQFLIAY-RQGADVLQNEGE 59
 QY 61 VTSHLQGMENFTIRLENGIKPVYVFDGKPPDMKKQELAKRYSKRDATKDCLEAVEVG 120
 DB 60 TTS-LMGMYRTMR-MENGIKPVYVFDGKPPQLKSGELAKSERAEKQJQQAQAGM 117
 QY 121 KDAIEKLSKRTVKTROHNEDECKRLRLMSVPVVEAPSEAECAALCINDKVAVASED 180
 DB 118 EEVEKEFTRLVKTAKYKHNDCKHLLSLMGIPYLDAPSEASQCALAKKGVAAATFD 177
 QY 181 MSLTEGAPRFLRLHMDPSSKKIPVMEFDVAKVLEELTMDQFIDLCILGCGCYCSIK 240
 DB 178 MDCLIFGSPVAKRHLIASEAKKLIQEFHLSRVLOELQNEQVDCILGSPCYCSIR 237
 QY 241 GIGGICLALKLROHGSLESLINKDKYQIPEDMFTYQEARCKEPPNVT-LDIPELKWT 299
 DB 238 GIGAKRAVDLIQKHISIEIVRRGDPKSYVPEXWLMHEAQCLPELEVVDPSEVFLKMS 297
 QY 300 APDEGLSFLVKDNGENEDRVTAIEIKSAKAKSSQGRLESPFKPATISALAKREE 359
 DB 298 EPNEBELVKTMGEGEQFEERIRISGVKRLSKRSGSTOGRLDDEFFKV7GSLSS-AKKEEP 356
 QY 360 SDKTSKAAANKKTKAGCKKK 379
 DB 357 F---PKGPAAKKKAKTGAAGK 373

RESULT 12

US-08-823-516-139
 Sequence 139, Application US/08823516
 Patent No. 5994869
 GENERAL INFORMATION:
 APPLICANT: Hall, Jeff G.
 APPLICANT: Lyons, Victor L.
 APPLICANT: Mast, Andrea L.
 APPLICANT: Bray, Mary Ann G.
 TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
 NUMBER OF SEQUENCES: 163
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Medien & Carroll, LLP
 STREET: 220 Montgomery Street, Suite 2200
 CITY: San Francisco
 STATE: California
 COUNTRY: United States Of America
 ZIP: 94104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk

US-08-823-516-79
 : Sequence 79, Application US/08823516
 : Patent No. 5994069
 : GENERAL INFORMATION:
 : APPLICANT: Hall, Jeff G.
 : APPLICANT: Lyamichev, Victor I.
 : APPLICANT: Mast, Mary Ann D.
 : APPLICANT: Brow, Mary Ann D.
 : TITLE OF INVENTION: Detection of Nucleic Acids By Multiple
 : NUMBER OF SEQUENCES: 163
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Medlen & Carroll, LLP
 : STREET: 220 Montgomery Street, Suite 2200
 : CITY: San Francisco
 : STATE: California
 : COUNTRY: United States of America
 : ZIP: 94104
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent In Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/823,516
 : FILING DATE: 24-MAR-1997
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: PCT/US97/01072
 : FILING DATE: 21-JAN-1997
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/759,038
 : FILING DATE: 02-DEC-1996
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/758,314
 : FILING DATE: 02-DEC-1996
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/756,386
 : FILING DATE: 29-NOV-1996
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/632,853
 : FILING DATE: 12-JUL-1996
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/599,491
 : FILING DATE: 24-JAN-1996
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Ingolia, Diane E.
 : REGISTRATION NUMBER: 40,027
 : REFERENCE/DOCKET NUMBER: FORS-02736
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (415) 705-8410
 : TELEFAX: (415) 397-8338
 : INFORMATION FOR SEQ ID NO: 79:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 340 amino acids
 : TYPE: amino acid
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : US-08-823-516-79

Query Match 33.4% Score 648.5; DB 2; Length 340;
 Best Local Similarity 41.3%; Pred. No. 7.4e-56;
 Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;

QY 19 KEQKESYFGKRAVADSMISYFLIVVGTGKETTINFAEYVTSKLCQGFNFTRPLFA 78
 DB 12 KEIIFENLYGKRIATLAINAYOFITIROKDGIPLMDSKGRITSLHSGLFVYITLMDRA 71
 QY 79 GIKPVYVFDGKPPDMKQSLAKRYSRDDATKDLTEAVSEVGDADIEKLSKRTVKTPOH 138
 DB 72 GIKPVYVFDGKPPDMKQSLAKRYSRDDATKDLTEAVSEVGDADIEKLSKRTVKTPOH 131
 QY 139 NEDCKRLRLMGVYVEAPSEAEACALCINDKVAVASECDMSLTFCAFRLRLMLKQ 198

US-08-823-516-79
 : Sequence 136, Application US/08823516
 : Patent No. 5994069
 : GENERAL INFORMATION:
 : APPLICANT: Hall, Jeff G.
 : APPLICANT: Lyamichev, Victor I.
 : APPLICANT: Mast, Mary Ann D.
 : APPLICANT: Brow, Mary Ann D.
 : TITLE OF INVENTION: Detection of Nucleic Acids By Multiple
 : NUMBER OF SEQUENCES: 163
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Medlen & Carroll, LLP
 : STREET: 220 Montgomery Street, Suite 2200
 : CITY: San Francisco
 : STATE: California
 : COUNTRY: United States of America
 : ZIP: 94104
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent In Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/823,516
 : FILING DATE: 24-MAR-1997
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: PCT/US97/01072
 : FILING DATE: 21-JAN-1997
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/759,038
 : FILING DATE: 02-DEC-1996
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/758,314
 : FILING DATE: 02-DEC-1996
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/756,386
 : FILING DATE: 29-NOV-1996
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/632,853
 : FILING DATE: 12-JUL-1996
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/599,491
 : FILING DATE: 24-JAN-1996
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Ingolia, Diane E.
 : REGISTRATION NUMBER: 40,027
 : REFERENCE/DOCKET NUMBER: FORS-02736
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (415) 705-8410
 : TELEFAX: (415) 397-8338
 : INFORMATION FOR SEQ ID NO: 136:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 340 amino acids
 : TYPE: amino acid

QY 245 QTALEIIPQHSIESITENLRD---RYQIHEDMVYQEARLKEPYNVTDIPELKTAP 3
 DB 248 KKALEIVH-----SKDPLAFQKQSVDEYAIKKEFLINPVT-DNYVCLVWRDP 2
 QY 302 DREGISLFVNDGNFEDRYTKALEIKISAKNKSQGRLESFFK 345
 DB 296 DREGISLFVNDGNFEDRYTKALEIKISAKNKSQGRLESFFK 339

RESULT 15
 US-08-823-516-79
 : Sequence 136, Application US/08823516
 : Patent No. 5994069
 : GENERAL INFORMATION:
 : APPLICANT: Hall, Jeff G.
 : APPLICANT: Lyamichev, Victor I.
 : APPLICANT: Mast, Mary Ann D.
 : APPLICANT: Brow, Mary Ann D.
 : TITLE OF INVENTION: Detection of Nucleic Acids By Multiple
 : NUMBER OF SEQUENCES: 163
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Medlen & Carroll, LLP
 : STREET: 220 Montgomery Street, Suite 2200
 : CITY: San Francisco
 : STATE: California
 : COUNTRY: United States of America
 : ZIP: 94104
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent In Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/823,516
 : FILING DATE: 24-MAR-1997
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: PCT/US97/01072
 : FILING DATE: 21-JAN-1997
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/759,038
 : FILING DATE: 02-DEC-1996
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/758,314
 : FILING DATE: 02-DEC-1996
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/756,386
 : FILING DATE: 29-NOV-1996
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/632,853
 : FILING DATE: 12-JUL-1996
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/599,491
 : FILING DATE: 24-JAN-1996
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Ingolia, Diane E.
 : REGISTRATION NUMBER: 40,027
 : REFERENCE/DOCKET NUMBER: FORS-02736
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (415) 705-8410
 : TELEFAX: (415) 397-8338
 : INFORMATION FOR SEQ ID NO: 136:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 340 amino acids
 : TYPE: amino acid

STRANDEDNESS: not relevant
 TOPOLOGY: not relevant
 MOLECULE TYPE: protein
 08-08-823-516-136

Query Match 33 43; Score 648 5; DB 2; Length 340;
 Best Local Similarity 41.3%; Prod No. 7,4e-56;
 Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;

```

QY      19 KEQKPESEYGRKTIADASMSYQFLIVGPTGMEETLTNEAGEVTSHQGMENPTIILEEA 78
DB      12 KETILENLGKRIADALNAIYQFLSTIPQKDGTPIMDSKGRITSHLSGLPTPTLNLMFA 71
QY      79 GIKPYVVEFGKPPDKKKOHLAKRYSKRDDATIKDLTEAVEGDKDAIFKLSKRTYKVTRQH 138
DB      72 GIKPYVVEFGKPPDKKKOHLAKRYSKRDDATIKDLTEAVEGDKDAIFKLSKRTYKVTRQH 131
QY      139 NEDCKRLRLMGVPVVAASEAECALCINDEKVAVASPFMDSLTPGAPRLRHMDP 198
DB      132 IEDAKKLELMGPTVQAPPSKGEAOAAVMAKGSVYASASQDYDLSLEGARLVKRLTII 191
QY      199 SSRKIP-----VMEFDVAKYLEEELTMDQFIDLCILCGDY-CDSIKGIGG 244
DB      192 GKRLPGKNVYVEIKPELILE---EVLKFKLTKRELIELATIVGIDYINPGCKKSISG 247
QY      245 QTALKLTIRQGSISILENLKND---RYQIPEDWPVQEARRLFKKPNVTIIPKCKWTAR 301
DB      248 KKALEIVRH-----SKDPLAKFOKOSDWDVAIKKEPFLNPVT-DNKNVWMDP 295
QY      302 DEEGGISFLVKNDFNEDPVTKAIEKIKSANNKSSQGRLESFPR 345
DB      296 DEEGILKFLCDEHDPSEERKNGLEIRLKKAIKSGKSTLESMPK 339

```

Search completed: October 22, 2003, 12:44:01
 Job time : 31 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

QM protein - protein search, using sw model

Run on: October 22, 2003, 12:40:26 : Search time 31 Seconds
(without alignments)
2047.341 Million cell updates/sec

Title: US-09-805-311-2

Sequence: 1 MKIKGLTKLLADNAPKAKE.....SDKISKAAANKTKAGCKKK 379

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 629382 seqs, 157660630 residues

Total number of hits satisfying chosen parameters: 629382

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 100%
Listing first 45 summaries

Database:

Published Applications-AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PC1_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PC105_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1939	100.0	379	9	US-09-805-311-2
2	1939	100.0	379	9	US-09-805-311-6
3	1933	99.7	379	9	US-09-805-311-4
4	1933	99.7	379	9	US-09-805-311-8
5	1032	53.2	380	9	US-09-805-311-10
6	1032	53.2	380	11	US-09-940-244-137
7	1032	53.2	380	12	US-10-290-386-137
8	1032	53.2	380	14	US-10-033-297-137
9	1014	52.3	378	11	US-09-940-244-138
10	1014	52.3	378	12	US-10-290-386-138
11	1014	52.3	378	14	US-10-033-297-138
12	938	48.4	382	11	US-09-940-244-139
13	938	48.4	382	12	US-10-290-386-139
14	938	48.4	382	14	US-10-033-297-139
15	648.5	33.4	340	10	US-09-777-430A-27

16	648.5	33.4	340	11	US-09-940-244-79	Sequence 79, f
17	648.5	33.4	340	11	US-09-940-244-136	Sequence 136, f
18	648.5	33.4	340	12	US-10-290-386-79	Sequence 79, f
19	648.5	33.4	340	12	US-10-290-386-136	Sequence 136, f
20	648.5	33.4	340	14	US-10-033-297-79	Sequence 79, f
21	648.5	33.4	340	14	US-10-033-297-136	Sequence 136, f
22	637	32.9	340	11	US-09-940-244-365	Sequence 365, f
23	632.5	32.6	340	11	US-09-940-244-357	Sequence 357, f
24	626	32.3	340	11	US-09-940-244-384	Sequence 384, f
25	624	32.2	346	11	US-09-940-244-379	Sequence 379, f
26	615	31.7	340	11	US-09-940-244-402	Sequence 402, f
27	615	31.7	351	11	US-09-940-244-337	Sequence 337, f
28	595	30.7	351	11	US-09-940-244-341	Sequence 341, f
29	582	30.0	328	11	US-09-940-244-398	Sequence 398, f
30	572	29.5	350	11	US-09-940-244-394	Sequence 394, f
31	569	29.1	351	11	US-09-940-244-361	Sequence 361, f
32	560.5	28.9	326	10	US-09-777-430A-28	Sequence 28, f
33	560.5	28.9	326	11	US-09-777-430A-75	Sequence 75, f
34	560.5	28.9	326	11	US-09-940-244-135	Sequence 135, f
35	560.5	28.9	326	12	US-10-290-386-75	Sequence 75, f
36	560.5	28.9	326	12	US-10-290-386-135	Sequence 135, f
37	560.5	28.9	326	14	US-10-033-297-75	Sequence 75, f
38	560.5	28.9	326	14	US-10-033-297-135	Sequence 135, f
39	560	28.0	356	11	US-09-940-244-389	Sequence 389, f
40	542	28.0	335	11	US-09-940-244-149	Sequence 149, f
41	541.5	27.9	881	11	US-09-864-636A-560	Sequence 560, f
42	541.5	27.9	881	11	US-09-864-636A-566	Sequence 566, f
43	541.5	27.9	881	11	US-09-864-636A-568	Sequence 568, f
44	541.5	27.9	881	12	US-10-084-839-560	Sequence 560, f
45	541.5	27.9	881	12	US-10-084-839-566	Sequence 566, f

ALIGNMENTS

RESULT 1	US-09-805-311-2	Sequence 2, Application US/09805311
1	Patent No. US20010034886A1	
2	GENERAL INFORMATION:	
3	APPLICANT: Mahajan, Pramod B.	
4	TITLE OF INVENTION: Rad2/FEN-1 Orthologues and Uses	
5	TITLE OF INVENTION: Theoret	
6	FILE REFERENCE: 09610	
7	CURRENT APPLICATION NUMBER: US/09/905,311	
8	CURRENT FILING DATE: 2001-03-13	
9	PRIOR APPLICATION NUMBER: 09/426,557	
10	PRIOR FILING DATE: 1999-10-22	
11	PRIOR APPLICATION NUMBER: 60/212,342	
12	PRIOR FILING DATE: 1998-12-15	
13	NUMBER OF SEQ ID NOS: 10	
14	SOFTWARE: FASTSEQ for Windows Version 1.0	
15	SEQ ID NO: 2	
16	LENGTH: 379	
17	TYPE: PRT	
18	ORGANISM: Zea mays	
19	US-09-805-311-2	
20	Query Match	100.0% Score 1939, DB 9, Length 379
21	Best Local Similarity	100.0% Prot. No. 17e-174
22	Matches 379: Conservative	0: Mismatches 0: Indels 0: Gaps
23	1	1
24	1	1
25	1	1
26	1	1
27	1	1
28	1	1
29	1	1
30	1	1
31	1	1
32	1	1
33	1	1
34	1	1
35	1	1
36	1	1
37	1	1
38	1	1
39	1	1
40	1	1
41	1	1
42	1	1
43	1	1
44	1	1
45	1	1


```

QY 181 MSLSFGAPRFLRLHMDPSSKRIIPWEDVAVKVELELIMDOFIDLCILGCGYCSIX 240
DB 181 MSLSFGAPRFLRLHMDPSSKRIIPWEDVAVKVELELIMDOFIDLCILGCGYCSIX 240
QY 241 GIGGOTALKLIROHGSISILENLKDRYQIPEDMPYOEARLKEPNVTLIDPELKWTA 300
DB 241 GIGGOTALKLIROHGSISILENLKDRYQIPEDMPYOEARLKEPNVTLIDPELKWTA 300
QY 301 PDEGLISFLVKGNGFNEDRYTKALEIKSAKNSSGGRLESFPRATTISAPLKRKETS 360
DB 301 PDEGLISFLVKGNGFNEDRYTKALEIKSAKNSSGGRLESFPRATTISAPLKRKETS 360
QY 361 DKTSKAANKRTKAGGKK 379
DB 361 DKTSKAANKRTKAGGKK 379

```

```

RESULT 2
US-09-805-311-6
Sequence 6, Application US/09805311
Patent No. US20010034886A1
GENERAL INFORMATION:
APPLICANT: Mahajan, Pramod B.
TITLE OF INVENTION: Rad2/FEN-1 Orthologues and Uses
FILE REFERENCE: 0961D
CURRENT APPLICATION NUMBER: US/09/805,311
PRIOR FILING DATE: 2001-03-13
PRIOR APPLICATION NUMBER: 09/426,557
PRIOR FILING DATE: 1999-10-22
PRIOR APPLICATION NUMBER: 60/112,332
PRIOR FILING DATE: 1998-12-15
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 379
TYPE: PRT
ORGANISM: Zea mays
US-09-805-311-6

```

```

Query Match          100.0% Score 1939, DB 9, Length 379;
Best Local Similarity 100.0% Pred. No. 1,70-174;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGKGLTKLADNAPKANKKQKFESEYFGKRIAVDASMSIYQFLVVGRTGNETLTNENGE 60
DB 1 MGKGLTKLADNAPKANKKQKFESEYFGKRIAVDASMSIYQFLVVGRTGNETLTNENGE 60
QY 61 VTSHLOGMENPTILLEGKIKPVYVFDGKPPDMKKQGLAKRYSKRDPAITKDLTEAVEYGD 120
DB 61 VTSHLOGMENPTILLEGKIKPVYVFDGKPPDMKKQGLAKRYSKRDPAITKDLTEAVEYGD 120
QY 121 KQATEKSKRTYKVRQHNEDCKRLRLKMGVYVAVPSEAEACALCINXKVAVASSED 180
DB 121 KQATEKSKRTYKVRQHNEDCKRLRLKMGVYVAVPSEAEACALCINXKVAVASSED 180
QY 181 MSLSFGAPRFLRLHMDPSSKRIIPWEDVAVKVELELIMDOFIDLCILGCGYCSIX 240
DB 181 MSLSFGAPRFLRLHMDPSSKRIIPWEDVAVKVELELIMDOFIDLCILGCGYCSIX 240
QY 241 GIGGOTALKLIROHGSISILENLKDRYQIPEDMPYOEARLKEPNVTLIDPELKWTA 300
DB 241 GIGGOTALKLIROHGSISILENLKDRYQIPEDMPYOEARLKEPNVTLIDPELKWTA 300
QY 301 PDEGLISFLVKGNGFNEDRYTKALEIKSAKNSSGGRLESFPRATTISAPLKRKETS 360
DB 301 PDEGLISFLVKGNGFNEDRYTKALEIKSAKNSSGGRLESFPRATTISAPLKRKETS 360
QY 361 DKTSKAANKRTKAGGKK 379
DB 361 DKTSKAANKRTKAGGKK 379

```

```

RESULT 3
US-09-805-311-4
Sequence 4, Application US/09805311
Patent No. US20010034886A1
GENERAL INFORMATION:
APPLICANT: Mahajan, Pramod B.
TITLE OF INVENTION: Rad2/FEN-1 Orthologues and Uses
FILE REFERENCE: 0961D
CURRENT APPLICATION NUMBER: US/09/805,311
PRIOR FILING DATE: 2001-03-13
PRIOR APPLICATION NUMBER: 09/426,557
PRIOR FILING DATE: 1999-10-22
PRIOR APPLICATION NUMBER: 60/112,332
PRIOR FILING DATE: 1998-12-15
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 379
TYPE: PRT
ORGANISM: Zea mays
US-09-805-311-4

```

```

Query Match          99.7% Score 1933, DB 9, Length 379;
Best Local Similarity 99.7% Pred. No. 6,20-174;
Matches 378; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGKGLTKLADNAPKANKKQKFESEYFGKRIAVDASMSIYQFLVVGRTGNETLTNENGE 60
DB 1 MGKGLTKLADNAPKANKKQKFESEYFGKRIAVDASMSIYQFLVVGRTGNETLTNENGE 60
QY 61 VTSHLOGMENPTILLEGKIKPVYVFDGKPPDMKKQGLAKRYSKRDPAITKDLTEAVEYGD 120
DB 61 VTSHLOGMENPTILLEGKIKPVYVFDGKPPDMKKQGLAKRYSKRDPAITKDLTEAVEYGD 120
QY 121 KQATEKSKRTYKVRQHNEDCKRLRLKMGVYVAVPSEAEACALCINXKVAVASSED 180
DB 121 KQATEKSKRTYKVRQHNEDCKRLRLKMGVYVAVPSEAEACALCINXKVAVASSED 180
QY 181 MSLSFGAPRFLRLHMDPSSKRIIPWEDVAVKVELELIMDOFIDLCILGCGYCSIX 240
DB 181 MSLSFGAPRFLRLHMDPSSKRIIPWEDVAVKVELELIMDOFIDLCILGCGYCSIX 240
QY 241 GIGGOTALKLIROHGSISILENLKDRYQIPEDMPYOEARLKEPNVTLIDPELKWTA 300
DB 241 GIGGOTALKLIROHGSISILENLKDRYQIPEDMPYOEARLKEPNVTLIDPELKWTA 300
QY 301 PDEGLISFLVKGNGFNEDRYTKALEIKSAKNSSGGRLESFPRATTISAPLKRKETS 360
DB 301 PDEGLISFLVKGNGFNEDRYTKALEIKSAKNSSGGRLESFPRATTISAPLKRKETS 360
QY 361 DKTSKAANKRTKAGGKK 379
DB 361 DKTSKAANKRTKAGGKK 379

```

SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 8
LENGTH: 379
TYPE: PRT
ORGANISM: Zea mays
US-09-805-311-8

Query Match
Best Local Similarity 99.7%; Pred. No. 6,2e+174;
Matches 378; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGKGLTKLADNAPKAKKEQFESEYGRKIAVDASMSIYQFLIVSGTGMETLTNEAGE 60
DB 1 MGKGLTKLADNAPKAKKEQFESEYGRKIAVDASMSIYQFLIVSGTGMETLTNEAGE 60
QY 61 VTSHLQGMFNRTIRLEAGIKPVYVFDGKPPDMKKQELAKRSKRDQATKDLTAVVEGD 120
DB 61 VTSHLQGMFNRTIRLEAGIKPVYVFDGKPPDMKKQELAKRSKRDQATKDLTAVVEGD 120
QY 121 KDAIEKLSKRTVKTROHNEDECKRLRLMGVYVAPSEAEACALCTNKKVAVASGD 180
DB 121 KDAIEKLSKRTVKTROHNEDECKRLRLMGVYVAPSEAEACALCTNKKVAVASGD 180
QY 181 KDSLTFGAPRFLRLHLDSSSKKIPVMEFDVAKVLEELTMDQFDLCTLCGDCYCESIK 240
DB 181 KDSLTFGAPRFLRLHLDSSSKKIPVMEFDVAKVLEELTMDQFDLCTLCGDCYCESIK 240
QY 241 GIGGOTALKLIRQHGSIESTLENKORVOIPEMPPYQEARLKEPVTLDTPELKMTA 300
DB 241 GIGGOTALKLIRQHGSIESTLENKORVOIPEMPPYQEARLKEPVTLDTPELKMTA 300
QY 301 PDEGLISFLVKNQFNEDRYTKALEIKISAKNKSQGRLESFEKPTATTSAPLRKETS 360
DB 301 PDEGLISFLVKNQFNEDRYTKALEIKISAKNKSQGRLESFEKPTATTSAPLRKETS 360
QY 361 DKTSKAANKKTAKGKKK 379
DB 361 DKTSKAANKKTAKGKKK 379

RESULT 5
US-09-805-311-10
Sequence 10, Application US/0980311
Patent No. US20010034886A1
GENERAL INFORMATION:
APPLICANT: Mahajan, Pramod B.
TITLE OF INVENTION: Rad2/FEN-1 Orthologues and Uses
FILE REFERENCE: 0961D
CURRENT APPLICATION NUMBER: US/09/805,311
CURRENT FILING DATE: 2001-03-13
PRIOR APPLICATION NUMBER: 09/426,557
PRIOR FILING DATE: 1999-10-22
PRIOR APPLICATION NUMBER: 60/112,332
PRIOR FILING DATE: 1998-12-15
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 10
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-09-805-311-10

Query Match
Best Local Similarity 53.2%; Score 1032; DB 9; Length 380;
Matches 206; Conservative 69; Mismatches 98; Indels 10; Gaps 6;

QY 1 MGKGLTKLADNAPKAKKEQFESEYGRKIAVDASMSIYQFLIVSGTGMETLTNEAGE 60
DB 1 MGKGLTKLADNAPKAKKEQFESEYGRKIAVDASMSIYQFLIVSGTGMETLTNEAGE 60
QY 61 VTSHLQGMFNRTIRLEAGIKPVYVFDGKPPDMKKQELAKRSKRDQATKDLTAVVEGD 120
DB 61 VTSHLQGMFNRTIRLEAGIKPVYVFDGKPPDMKKQELAKRSKRDQATKDLTAVVEGD 120

DB 60 TISHLMKQEVRTIRKMNCKIKPVYVFDGKPPDLSQGLAKRSERAEAKKQDQQAAGA
QY 121 KDAIEKLSKRTVKTROHNEDECKRLRLMGVYVAPSEAEACALCTNKKVAVASGD 180
DB 121 KDAIEKLSKRTVKTROHNEDECKRLRLMGVYVAPSEAEACALCTNKKVAVASGD 180
QY 181 KDSLTFGAPRFLRLHLDSSSKKIPVMEFDVAKVLEELTMDQFDLCTLCGDCYCESIK 240
DB 181 KDSLTFGAPRFLRLHLDSSSKKIPVMEFDVAKVLEELTMDQFDLCTLCGDCYCESIK 240
QY 241 GIGGOTALKLIRQHGSIESTLENKORVOIPEMPPYQEARLKEPVTLDTPELKMTA 300
DB 241 GIGGOTALKLIRQHGSIESTLENKORVOIPEMPPYQEARLKEPVTLDTPELKMTA 300
QY 301 PDEGLISFLVKNQFNEDRYTKALEIKISAKNKSQGRLESFEKPTATTSAPLRKETS 360
DB 301 PDEGLISFLVKNQFNEDRYTKALEIKISAKNKSQGRLESFEKPTATTSAPLRKETS 360
QY 361 DKTSKAANKKTAKGKKK 379
DB 361 DKTSKAANKKTAKGKKK 379

RESULT 6
US-09-940-244-137
Sequence 137, Application US/09940244
Publication No. US20040044796A1
GENERAL INFORMATION:
APPLICANT: Netti, Bruce P.
APPLICANT: Hall, Jeff G.
APPLICANT: Lyamichiev, Victor
APPLICANT: Smith, Lloyd M.
TITLE OF INVENTION: Reactions on Dendrimers
FILE REFERENCE: PORS-06476
CURRENT APPLICATION NUMBER: US/09/940,244
CURRENT FILING DATE: 2002-05-06
NUMBER OF SEQ ID NOS: 422
SOFTWARE: patentin version 3.1
SEQ ID NO 137
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-09-940-244-137

Query Match
Best Local Similarity 53.8%; Score 1072; DB 11; Length 380;
Matches 206; Conservative 69; Mismatches 98; Indels 10; Gaps 6;

QY 1 MGKGLTKLADNAPKAKKEQFESEYGRKIAVDASMSIYQFLIVSGTGMETLTNEAGE 60
DB 1 MGKGLTKLADNAPKAKKEQFESEYGRKIAVDASMSIYQFLIVSGTGMETLTNEAGE 60
QY 61 VTSHLQGMFNRTIRLEAGIKPVYVFDGKPPDMKKQELAKRSKRDQATKDLTAVVEGD 120
DB 61 VTSHLQGMFNRTIRLEAGIKPVYVFDGKPPDMKKQELAKRSKRDQATKDLTAVVEGD 120
QY 121 KDAIEKLSKRTVKTROHNEDECKRLRLMGVYVAPSEAEACALCTNKKVAVASGD 180
DB 121 KDAIEKLSKRTVKTROHNEDECKRLRLMGVYVAPSEAEACALCTNKKVAVASGD 180
QY 181 KDSLTFGAPRFLRLHLDSSSKKIPVMEFDVAKVLEELTMDQFDLCTLCGDCYCESIK 240
DB 181 KDSLTFGAPRFLRLHLDSSSKKIPVMEFDVAKVLEELTMDQFDLCTLCGDCYCESIK 240
QY 241 GIGGOTALKLIRQHGSIESTLENKORVOIPEMPPYQEARLKEPVTLDTPELKMTA 300
DB 241 GIGGOTALKLIRQHGSIESTLENKORVOIPEMPPYQEARLKEPVTLDTPELKMTA 300
QY 301 PDEGLISFLVKNQFNEDRYTKALEIKISAKNKSQGRLESFEKPTATTSAPLRKETS 360
DB 301 PDEGLISFLVKNQFNEDRYTKALEIKISAKNKSQGRLESFEKPTATTSAPLRKETS 360
QY 361 DKTSKAANKKTAKGKKK 379
DB 361 DKTSKAANKKTAKGKKK 379

Db 358 PEPKSTKKKA---KTGAAGKFK 377

RESULT 7

US-10-290-386-137
 : Sequence 137, Application US/10290386
 : Publication No. US20030152971A1
 : GENERAL INFORMATION:
 : APPLICANT: Lyamichev, Victor
 : APPLICANT: Neri, Bruce P.
 : APPLICANT: Hall, Jeff G.
 : APPLICANT: Lukovlak, Andrew A.
 : TITLE OF INVENTION: Methods and Compositions for Detecting Target Sequences
 : FILE REFERENCE: FORS-07459
 : CURRENT APPLICATION NUMBER: US/10/290,386
 : PRIORITY FILING DATE: 2002-11-07
 : PRIORITY FILING DATE: 2002-02-27
 : PRIORITY FILING DATE: 2002-02-27
 : PRIORITY FILING DATE: 2001-11-07
 : PRIORITY FILING DATE: 2001-11-07
 : PRIORITY FILING DATE: 2000-11-15
 : PRIORITY FILING DATE: 2000-11-15
 : PRIORITY FILING DATE: 2000-02-08
 : PRIORITY FILING DATE: 2000-02-08
 : PRIORITY FILING DATE: 1999-07-09
 : PRIORITY FILING DATE: 1999-07-09
 : PRIORITY FILING DATE: 1997-03-24
 : PRIORITY FILING DATE: 1997-03-24
 : PRIORITY FILING DATE: 1996-12-02
 : PRIORITY FILING DATE: 1996-12-02
 : PRIORITY FILING DATE: 1996-11-26
 : PRIORITY FILING DATE: 1996-07-12
 : PRIORITY FILING DATE: 1996-07-12
 : PRIORITY FILING DATE: 08/599,491
 : NUMBER OF SEQ ID NOS: 253
 : SOFTWARE: PatentIn version 3.1
 : SEQ ID NO 137
 : LENGTH: 380
 : TYPE: PRT
 : ORGANISM: Artificial
 : FEATURE:
 : OTHER INFORMATION: Synthetic
 : US-10-290-386-137

Query Match 53.2% Score 1032 DB 22 Length 380
 Best Local Similarity 53.8% Pred. No. 7.2e-89
 Matches 206 Conservative 69 Mismatches 98 Indels 10 Gaps 6

1 MGIGKGLTKRIADNAPAMKRFKFFSGCRKIVADASMSITQFLLTVGCRGMRLLTNEAGF 60
 1 MGTGGLAKLIADNAPAMKRFKFFSGCRKIVADASMSITQFLLTVGCRGMRLLTNEAGF 59
 61 VTSILQGMFNRTRJLEAGTRPVYVDFGKPPDKMKOELAKRYSKSDATKDLTFAVAVG 120
 60 TTSILQGMFNRTRJLEAGTRPVYVDFGKPPDKMKOELAKRYSKSDATKDLTFAVAVG 119
 121 KDAIEKLSKRTVAVTQHNDCRRLRLKAGVPVVEAPFAEACALCINDKYFAVASED 160
 120 EQEVEKTRKLVAVTQHNDCRRLRLKAGVPVVEAPFAEACALCINDKYFAVASED 179
 181 MSLTFGAPRFLRHLMDPSSKRPVMEFFDAKYLELELTMDQFIDLICGCGDYGDSTK 240
 180 MDCITFGAPRFLRHLMDPSSKRPVMEFFDAKYLELELTMDQFIDLICGCGDYGDSTK 239
 241 GIGGQALAKLIRHGSSTSLNKLKDRQICPRDWPYCEARRLKPEPNVTL--IPELAK 298
 240 GIGGQALAKLIRHGSSTSLNKLKDRQICPRDWPYCEARRLKPEPNVTL--IPELAK 298
 299 TAPDEGLISFLXKDGQFNEDRVTKALEIKSAKNNSSQGRLESFFKPRNTSAPLKRF 358
 298 TAPDEGLISFLXKDGQFNEDRVTKALEIKSAKNNSSQGRLESFFKPRNTSAPLKRF 358

Db 299 SEPNEEELCKEKGCKQSEERIRSGVRLSKSRQSTQSRIDDFPKYTSLSL-ARKE 1
 Cy 359 TSDK---TSKAAANKKTKAGCKK 379
 Db 358 PEPKSTKKKA---KTGAAGKFK 377

RESULT 8

US-10-033-297-137
 : Sequence 137, Application US/10033297
 : Publication No. US20020187486A1
 : GENERAL INFORMATION:
 : APPLICANT: Hall, Jeff G.
 : APPLICANT: Lyamichev, Victor I.
 : APPLICANT: Mast, Andrea L.
 : APPLICANT: Brown, Mary Ann P.
 : TITLE OF INVENTION: Detection of Nucleic Acids By Multiple
 : CORRESPONDENCE ADDRESS:
 : ADDRESS: Medion & Carroll, LLP
 : STREET: 220 Montgomery Street, Suite 2200
 : CITY: San Francisco
 : STATE: California
 : COUNTRY: United States of America
 : ZIP: 94104
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/10/033,297
 : FILING DATE: 12-NOV-2002
 : CLASSIFICATION: C04B0001
 : PRIORITY APPLICATION DATA:
 : APPLICATION NUMBER: US/09/350,597
 : FILING DATE: 09-JUL-1999
 : APPLICATION NUMBER: US/08/823,515
 : FILING DATE: 24-MAR-1997
 : APPLICATION NUMBER: PCT/US97/01072
 : FILING DATE: 21-JAN-1997
 : APPLICATION NUMBER: US 08/759,038
 : FILING DATE: 02-DEC-1996
 : APPLICATION NUMBER: US 08/756,314
 : FILING DATE: 02-DEC-1996
 : APPLICATION NUMBER: US 08/756,386
 : FILING DATE: 29-NOV-1996
 : APPLICATION NUMBER: US 08/682,853
 : FILING DATE: 12-JUL-1996
 : APPLICATION NUMBER: US 08/599,491
 : FILING DATE: 28-JAN-1996
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Foglia, David B.
 : REGISTRATION NUMBER: 40,027
 : REFERENCE/DOCKET NUMBER: FORS-02736
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (415) 705-8410
 : TELEFAX: (415) 397-8338
 : INFORMATION FOR SEQ ID NO: 137:
 : LENGTH: 360 amino acids
 : TYPE: amino acid
 : STRANDEDNESS: No
 : MOLECULE TYPE: Protein
 : TOPOLOGY: No
 : SEQUENCE DESCRIPTION: SEQ ID NO: 137:
 : US-10-033-297-137

Query Match 53.2% Score 1032 DB 14 Length 380
 Best Local Similarity 53.8% Pred. No. 7.2e-92
 Matches 206 Conservative 69 Mismatches 98 Indels 10 Gaps 6

```

QY 1 MGICG:IKLLADNAPKAMEQKFESEYFGKIAVDASMSIYOFLLIYVGRFZMETLTNPAGE 60
DB 1 MGICGLAKLADVAPSAITENDIKSYFGKVAIDASMSIYOFLLIYVGRFZMETLTNPAGE 59
QY 61 VTSILOGMENRTIRLEAGIKPVYVFDGKPPDMKQOELAKRYSKROCATKOLTEAVEYGD 120
DB 60 TSHLMGMFYRTIRWMENGIKPVYVFDGKPPOLKSGELAKRSEBAEAKQLOQOAGAGA 119
QY 121 KDAIEKLSKRTYKVRQHINDECKRLILNGVYVVAAPSAEACALCINDKVFVAASED 180
DB 120 EQVEVEKTRIAVKVRKQHNDECKRLILSLMGIPYIDAPSEAEKALVAKRKYAAALIEC 179
QY 181 MDLTFGAPRFLRHIMDPSKRIPIWEPYVAVLEBELTMDQFCTLTIGCGVYVTSIRK 240
DB 180 MDLTFGAPRFLRHIMDPSKRIPIWEPYVAVLEBELTMDQFCTLTIGCGVYVTSIRK 239
QY 241 GIGOGTALKLIRQHSIESILENLDKRYOIEEDAPYQBARLSEKPNVTJD-1PECKK 296
DB 240 GIGOPRAVDLQKHKSEIEIVRLDPNKKYPDENMLKHAQILFLEPEV-LDPSEVSEKWM 298
QY 299 TAPDEGLISFLVCKGFMEDRVTKAIKIKSAKKSSQGLSESPFKRTATISAPLKKE 358
DB 298 SEPNEELIKFMCGEKOFSEERIRSGVSKLSKROGSOGRLDFFKVGSLSS-AKRKE 357
QY 359 TSDK-TSKAANKTKRACGKK 379
DB 358 PEPKSTKKKA---KTGAAGKFK 377

```

```

RESULT 9
US-09-940-244-138
Sequence 138, Application US/09940244
Publication No. US20030044796A1
GENERAL INFORMATION:
APPLICANT: Neti, Bruce P.
APPLICANT: Hall, Jeff G.
APPLICANT: Lyamichev, Victor
APPLICANT: Smith, Lloyd M.
TITLE OF INVENTION: Reactions on dendrimers
FILE REFERENCE: FORS-06478
CURRENT APPLICATION NUMBER: US/09/940,244
CURRENT FILING DATE: 2002-05-06
NUMBER OF SEQ ID NOS: 422
SOFTWARE: PatentIn version 3.1
SEQ ID NO 138
LENGTH: 378
TYPE: PRI
ORGANISM: Mus musculus
US-09-940-244-138

```

```

Query Match 52.3%; Score 1014; DB 11; Length 378;
Best Local Similarity 53.2%; Pred. No. 3,6e-87;
Matches 202; Conservative 59; Mismatches 101; Indels 8; Gaps 6;

```

```

QY 1 MGICG:IKLLADNAPKAMEQKFESEYFGKIAVDASMSIYOFLLIYVGRFZMETLTNPAGE 60
DB 1 MGICGLAKLADVAPSAITENDIKSYFGKVAIDASMSIYOFLLIYVGRFZMETLTNPAGE 59
QY 61 VTSILOGMENRTIRLEAGIKPVYVFDGKPPDMKQOELAKRYSKROCATKOLTEAVEYGD 120
DB 60 TSHLMGMFYRTIRWMENGIKPVYVFDGKPPOLKSGELAKRSEBAEAKQLOQOAGAGA 119
QY 121 KDAIEKLSKRTYKVRQHINDECKRLILNGVYVVAAPSAEACALCINDKVFVAASED 180
DB 120 EQVEVEKTRIAVKVRKQHNDECKRLILSLMGIPYIDAPSEAEKALVAKRKYAAALIEC 179
QY 181 MDLTFGAPRFLRHIMDPSKRIPIWEPYVAVLEBELTMDQFCTLTIGCGVYVTSIRK 240
DB 180 MDLTFGAPRFLRHIMDPSKRIPIWEPYVAVLEBELTMDQFCTLTIGCGVYVTSIRK 239
QY 241 GIGOGTALKLIRQHSIESILENLDKRYOIEEDAPYQBARLSEKPNVTJD-1PECKK 296
DB 240 GIGOPRAVDLQKHKSEIEIVRLDPNKKYPDENMLKHAQILFLEPEV-LDPSEVSEKWM 298
QY 299 TAPDEGLISFLVCKGFMEDRVTKAIKIKSAKKSSQGLSESPFKRTATISAPLKKE 358
DB 298 SEPNEELIKFMCGEKOFSEERIRSGVSKLSKROGSOGRLDFFKVGSLSS-AKRKE 357

```

```

QY 300 APDEGLISFLVCKGFMEDRVTKAIKIKSAKKSSQGLSESPFKRTATISAPLKKE 358
DB 298 EPNEELIKFMCGEKOFSEERIRSGVSKLSKROGSOGRLDFFKVGSLSS-AKRKE 357
QY 357 E---PKSIKAKKATGAGK 373
DB 357 E---PKSIKAKKATGAGK 373

```

```

RESULT 10
US-10-290-386-138
Sequence 138, Application US/10290386
Publication No. US2003015297A1
GENERAL INFORMATION:
APPLICANT: Neti, Bruce P.
APPLICANT: Hall, Jeff G.
APPLICANT: Lyamichev, Victor
APPLICANT: Luskowak, Andrew A.
TITLE OF INVENTION: Methods and Compositions for Detecting Target Sequences
FILE REFERENCE: FORS-07459
CURRENT APPLICATION NUMBER: US/10/290,386
CURRENT FILING DATE: 2002-11-07
PRIOR APPLICATION NUMBER: 60/351,060
PRIOR FILING DATE: 2002-02-27
PRIOR APPLICATION NUMBER: 60/344,946
PRIOR FILING DATE: 2001-11-07
PRIOR APPLICATION NUMBER: 69/7713,602
PRIOR FILING DATE: 2000-11-15
PRIOR APPLICATION NUMBER: 09/381,212
PRIOR FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: 09/350,335
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 08/823,516
PRIOR FILING DATE: 1997-03-24
PRIOR APPLICATION NUMBER: 08/759,018
PRIOR FILING DATE: 1996-12-02
PRIOR APPLICATION NUMBER: 08/756,386
PRIOR FILING DATE: 1996-11-26
PRIOR APPLICATION NUMBER: 08/662,853
PRIOR FILING DATE: 1996-07-12
PRIOR APPLICATION NUMBER: 08/599,491
PRIOR FILING DATE: 1996-01-24
NUMBER OF SEQ ID NOS: 253
SOFTWARE: PatentIn version 3.1
SEQ ID NO 138
LENGTH: 378
TYPE: PRI
ORGANISM: Artificial
FEATURES:
OTHER INFORMATION: Synthetic
US-10-290-386-138

```

```

Query Match 52.3%; Score 1014; DB 12; Length 378;
Best Local Similarity 53.2%; Pred. No. 3,6e-87;
Matches 202; Conservative 59; Mismatches 101; Indels 8; Gaps 6;

```

```

QY 1 MGICG:IKLLADNAPKAMEQKFESEYFGKIAVDASMSIYOFLLIYVGRFZMETLTNPAGE 60
DB 1 MGICGLAKLADVAPSAITENDIKSYFGKVAIDASMSIYOFLLIYVGRFZMETLTNPAGE 59
QY 61 VTSILOGMENRTIRLEAGIKPVYVFDGKPPDMKQOELAKRYSKROCATKOLTEAVEYGD 120
DB 60 TSHLMGMFYRTIRWMENGIKPVYVFDGKPPOLKSGELAKRSEBAEAKQLOQOAGAGA 119
QY 121 KDAIEKLSKRTYKVRQHINDECKRLILNGVYVVAAPSAEACALCINDKVFVAASED 180
DB 120 EQVEVEKTRIAVKVRKQHNDECKRLILSLMGIPYIDAPSEAEKALVAKRKYAAALIEC 179
QY 181 MDLTFGAPRFLRHIMDPSKRIPIWEPYVAVLEBELTMDQFCTLTIGCGVYVTSIRK 240
DB 180 MDLTFGAPRFLRHIMDPSKRIPIWEPYVAVLEBELTMDQFCTLTIGCGVYVTSIRK 239
QY 241 GIGOGTALKLIRQHSIESILENLDKRYOIEEDAPYQBARLSEKPNVTJD-1PECKK 296
DB 240 GIGOPRAVDLQKHKSEIEIVRLDPNKKYPDENMLKHAQILFLEPEV-LDPSEVSEKWM 298
QY 299 TAPDEGLISFLVCKGFMEDRVTKAIKIKSAKKSSQGLSESPFKRTATISAPLKKE 358
DB 298 SEPNEELIKFMCGEKOFSEERIRSGVSKLSKROGSOGRLDFFKVGSLSS-AKRKE 357

```


Db 178 MOTLCVPTPLRHLTFSSAKKEPIHEIDPEIWLRLGDLTLEOFVDCIMLGDCYCESTR 237
 QY 241 GIGGOTALKLIRHOSHIESLENNL-----NKDRYOIPEDMPYQEARLFEKPNVLDLPE 235
 Db 238 GVGPPVATLKLIKIHGSHIEKVEFIESGESNNTKMKIPEDMPYQAKRLELDPEV-IDGNE 236
 QY 296 --LKWTAPEDEGLISFLVKNQNGENEDRVTKALEIKSAKNSQGRLESFFKPTATT--- 240
 Db 297 INLKWSPKREKELIPEYICDDKKFSERVKSGISRLKGLKSGIOGRDLDFQVVFMTKEQ 236
 QY 351 -SAPLKKRETSOKTSAANKKTK 373
 Db 357 LAAAKKROENKRLNK-NKNKVTK 379

RESULT 13
 US-10-290-366-139
 : Sequence 139, Application US/10290386
 : Publication No. US20030152971A1
 : GENERAL INFORMATION:
 : APPLICANT: Lyamichev, Victor
 : APPLICANT: Hall, Jeff G.
 : APPLICANT: Lakowatz, Andrew A.
 : TITLE OF INVENTION: Methods and Compositions for Detecting Target Sequences
 : FILE REFERENCE: FORS-07459
 : CURRENT APPLICATION NUMBER: US/10/290,386
 : PRIOR FILING DATE: 2002-11-07
 : PRIOR APPLICATION NUMBER: 60/361,060
 : PRIOR FILING DATE: 2002-02-27
 : PRIOR APPLICATION NUMBER: 60/344,946
 : PRIOR FILING DATE: 2001-11-07
 : PRIOR APPLICATION NUMBER: 09/713,601
 : PRIOR FILING DATE: 2000-11-15
 : PRIOR APPLICATION NUMBER: 08/381,212
 : PRIOR FILING DATE: 2000-02-08
 : PRIOR APPLICATION NUMBER: 09/350,309
 : PRIOR FILING DATE: 1999-07-09
 : PRIOR APPLICATION NUMBER: 08/823,516
 : PRIOR FILING DATE: 1997-03-24
 : PRIOR APPLICATION NUMBER: 08/759,038
 : PRIOR FILING DATE: 1996-12-02
 : PRIOR APPLICATION NUMBER: 08/756,386
 : PRIOR FILING DATE: 1996-11-26
 : PRIOR APPLICATION NUMBER: 08/682,853
 : PRIOR FILING DATE: 1996-07-12
 : PRIOR APPLICATION NUMBER: 08/559,491
 : PRIOR FILING DATE: 1996-01-24
 : NUMBER OF SEQ ID NOS: 253
 : SOFTWARE: Patent version 3.1
 : SEQ ID NO 139
 : LENGTH: 362
 : TYPE: PRT
 : ORGANISM: Artificial
 : FEATURE:
 : OTHER INFORMATION: Synthetic
 : US-10-290-366-139

Query Match 48.4%; Score 938; DB 12; Length 362;
 Best Local Similarity 49.5%; Pred. No. 5,5e-60;
 Matches 190; Conservative 69; Mismatches 109; Indels 16; Gaps 6;

QY 1 MG-KGLTKLADNAPKAKKQKESFFGRKIVADASYSTIQLIVVGRMGTLLTNADE 60
 Db 1 MGKIGLNATLSEHVAFSAIKSDIKSEGRKVALIDASKSCQFLIVARQOGRGLTNADE 60
 QY 61 VSHLQGMENRTIRLLEAGIKPYVVDKPPDKKQDELAKRYSKQDARIDLTVANEVQ 120
 Db 61 TSHLQGMENRTIRLLEAGIKPYVVDKPPDKKQDELAKRYSKQDARIDLTVANEVQ 120
 QY 121 KDAIEKSKRTVAVTRQHNEDCKRLRLIMGVPYVAEPAEAEACALCTINDKVFVASE 180
 Db 121 KDAIEKSKRTVAVTRQHNEDCKRLRLIMGVPYVAEPAEAEACALCTINDKVFVASE 180

Db 121 K---MKQBRVLVKSNEENEAOKLJLGMGIPYTAPEAEOCAELAKKGVYAAASED 1
 QY 181 MDSITGAPRLAH-MUPSSKRIPWKEFVAVVLELELTMOQFTDCTLGGCCVCDSTK 2
 Db 178 KOTLCVPTPLRHLTFSSAKKEPIHEIDPEIWLRLGDLTLEOFVDCIMLGDCYCESTR 237
 QY 241 GIGGOTALKLIRHOSHIESLENNL-----NKDRYOIPEDMPYQEARLFEKPNVLDLPE 235
 Db 238 GVGPPVATLKLIKIHGSHIEKVEFIESGESNNTKMKIPEDMPYQAKRLELDPEV-IDGNE 236
 QY 296 --LKWTAPEDEGLISFLVKNQNGENEDRVTKALEIKSAKNSQGRLESFFKPTATT--- 240
 Db 297 INLKWSPKREKELIPEYICDDKKFSERVKSGISRLKGLKSGIOGRDLDFQVVFMTKEQ 236
 QY 351 -SAPLKKRETSOKTSAANKKTK 373
 Db 357 LAAAKKROENKRLNK-NKNKVTK 379

RESULT 14
 US-10-033-297-139
 : Sequence 139, Application US/10033297
 : Publication No. US20020187486A1
 : GENERAL INFORMATION:
 : APPLICANT: Hall, Jeff G.
 : APPLICANT: Lyamichev, Victor I.
 : APPLICANT: Brown, Mary Ann D.
 : TITLE OF INVENTION: Detection of Nucleic Acids By Multiple
 : Sequential Invasive Cleavages
 : NUMBER OF SEQUENCES: 163
 : CORRESPONDENCE ADDRESS:
 : ADDRESS: Medien & Carroll, LLP
 : STREET: 220 Montgomery Street, Suite 2200
 : CITY: San Francisco
 : STATE: California
 : COUNTRY: United States of America
 : ZIP: 94104
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: floppy disk
 : OPERATING SYSTEM: PC/POS/MS-DOS
 : SOFTWARE: Patent in Re-Case #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/10/033,297
 : FILING DATE: 12-NOV-02US20020187486A1-2001
 : CLASSIFICATION: <UNKNOWN>
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US/09/350,597
 : FILING DATE: 09-JUL-1999
 : APPLICATION NUMBER: US/04/823,516
 : FILING DATE: 24-MAR-1997
 : APPLICATION NUMBER: PCT/US97/01072
 : FILING DATE: 21-JAN-1997
 : APPLICATION NUMBER: US 06/759,038
 : FILING DATE: 02-DEC-1996
 : APPLICATION NUMBER: US 04/758,314
 : FILING DATE: 02-DEC-1996
 : APPLICATION NUMBER: US 08/756,186
 : FILING DATE: 29-NOV-1996
 : APPLICATION NUMBER: US 06/682,853
 : FILING DATE: 12-JUL-1996
 : APPLICATION NUMBER: US 08/559,491
 : FILING DATE: 24-JAN-1996
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Lyamichev, Victor I.
 : REGISTRATION NUMBER: 40,027
 : REFERENCE/DOCKET NUMBER: FORS-02736
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (415) 705-8410
 : TELEFAX: (415) 397-8338
 : INFORMATION FOR SEQ ID NO: 139:
 : SEQUENCE CHARACTERISTICS:

LENGTH: 382 amino acids
 TYPE: amino acid
 STRANDEDNESS: NO, US20020187486A1 Relevant
 TOPOLOGY: NO, US20020187486A1 Relevant
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 139
 US-10-033-297-139

Query Match 48.4% Score 938; DB 14; Length 382;
 Best Local Similarity 49.5%; Pred. No. 5,56-80;
 Matches 190; Conservative 69; Mismatches 105; Indels 16; Gaps 6;

QY 1 MGICGLTLADNAKAKKFKFESYFGRKLAIVASXSIYQFLVVGRTGSETLTKERGE 60
 DB 1 MGICGLTLADNAKAKKFKFESYFGRKLAIVASXSIYQFLVVGRTGSETLTKERGE 60
 QY 6 VTSEJGMPNRTIRLEAGIKPVYVFDGKPPDMKQGLAKRYSKDDTKDLTLEAVETG 120
 DB 6 VTSEJGMPNRTIRLEAGIKPVYVFDGKPPDMKQGLAKRYSKDDTKDLTLEAVETG 120
 QY 121 KDAIEKLSKRTVYRQHNEDSKRLRLMGVYVVAPEAEAFACALCINKEVAVASED 180
 DB 121 K--MKGERRLVYKSGHNEEOKLLGKGIYIIAPTEFAAGCALAKKRVAAASRD 177
 QY 181 MSLITGAPRFLRLMDFSSKKIPVMEFDVAKVEDELIMDQFLDLCIIGGQDYCSIK 240
 DB 178 MDTLCYRTPFLRLHLESEAKKEPHEIDTDLVGLDLTECFVDLCTMLGQDYCSIR 237
 QY 241 GIGGQTALKLIRHGSISILENC-----SKDRQIPEDNPQGRK--FKENVTLDIPE 285
 DB 238 GVGPTALKLIKTRHSITKIVFIEISGESNNTKAIPEQMPYKQAKMLJLDPNV--IDENE 286
 QY 296 --LKWTADEGLISFLVKGNGFNEDRYTKALEIKTSAKNKSQGRLESEFFKPA--ATT--- 350
 DB 297 TLKWSPPREKELYLDODKKSESEYKSGISRLKGLKSGIQRLDGLFFGVVFKTKEQ 356
 QY 351 -SAPLRKKTSDKTSKAANKTK 373
 DB 357 IAAAKRAQENKLNK--KNKVTK 379

RESULT 15

US-09-777-430A-27
 ? Sequence 27, Application US/C9777430A
 ? Patent No. US20020128465A1
 ? GENERAL INFORMATION:
 ? APPLICANT: Lyamichev, Victor
 ? APPLICANT: Skrzypczynski, Zbigniew
 ? APPLICANT: Allawi, Hatim T.
 ? APPLICANT: Mayland, Sarah R.
 ? APPLICANT: Yakova, Tsetska
 ? APPLICANT: Neil, Bruce P.
 ? TITLE OF INVENTION: Charge Tags and the Separation of Nucleic Acid Molecules
 ? FILE REFERENCE: FORS-04912
 ? CURRENT APPLICATION NUMBER: US/09/777,430A
 ? CURRENT FILING DATE: 2001-02-06
 ? NUMBER OF SEQ ID NOS: 85
 ? SOFTWARE: PatentIn version 3.1
 ? SEQ ID NO 27
 ? LENGTH: 340
 ? TYPE: PRT
 ? ORGANISM: *Pyrococcus furiosus*
 US-09-777-430A-27

Query Match 33.4% Score 648.5; DB 10; Length 340;
 Host Local Similarity 41.3%; Pred. No. 9,9e-53;
 Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 0;

QY 19 KFOKESTFCRKLAVDAKMSITVQELIVGRTRMFTLINAGVYVSHCGMFRNRTPLTFA 74
 DB 12 KELEKNTLGRKIALDALNATVQFLSTIRQKIDGTPCLMDSKGLTSHLSGLPYRITINDEA 71
 QY 79 GIKPVYVFDGKPPDMKQGLAKRYSKDDTKDLTLEAVETGDKAIETKISRIYKVIKSH 138

DB 72 GIKPVYVFDGKPPDMKQGLAKRYSKDDTKDLTLEAVETGDKAIETKISRIYKVIKSH 138
 QY 139 NEDCKPLLSMGVYVVAPEAEAE--ALGCTINDKVFVAVASEDMDSLTFCGAPRFLRHMDP 1
 DB 132 IFDAKKULELMGIPYVQAPSEGEAOAAVMAKGSYVASASQDUSLTFCAPIRLVRLNIT 1
 QY 199 SSKKIP-----VMEFDVAKVEDELIMDQFLDLCIIGGQDY--CSSTKGTG 2
 DB 192 GRRKLPGRKNYVVELKPR--ILR-----EVLKELKTFSEKLELA--IVGTQVNPQIKIGL 2
 QY 245 CTALKLIRHGSISILENC-----RYQIPETMPYQEARMLKEHNVITLDIPLKATAP 3
 DB 248 KKALEIVRH-----SKDPLAKFGQGVYDLYAIKKEFFLNPVI--FNQNVLMWRD 2
 QY 302 LEEELISFLVKGNGFNEDRYTKALEIKTSAKNKSQGRLESEFFK 345
 DB 296 DEKDLKFLDDEHDSERKYSGERLKKAIKSKQSTIDSMER 339

Search completed: October 22, 2003, 12:44:18
 Job time : 32 secs